

**CLONING AND CHARACTERIZATION OF THE UBIQUITIN
CARBOXYL TERMINAL HYDROLASE GENE IN *DROSOPHILA*
*MELANOGASTER***

NIAN ZHANG

Thesis presented for the Degree of Doctor of Philosophy
University of Edinburgh
1992



DEDICATION

To my dear wife Liang Kang for all the suffers she has had to support my doing this degree.

DECLARATION

I declare that this thesis was composed by myself, and that the work described is my own, unless otherwise stated.

Nian Zhang
February, 1992

ABSTRACT

In this project, 61 strains of *Drosophila melanogaster* carrying the enhancer detector P[lac, ry⁺]A were stained for β -galactosidase activity in the adult gonad. This P-transposon detects regulatory elements of a nearby gene. Sex- and tissue-specific staining were found in many strains. Some strains showed staining in very specific cells. These cells may play very important roles during oogenesis.

16.5 kb genomic fragment surrounding the enhancer detector P[lac, ry⁺], located at 98F of chromosome 3, has been cloned from a strain called D19. This strain shows the *lacZ* staining in germ-line cells of the adult ovary and testis. This DNA fragment contains multiple transcription units. A 1.3 kb HindIII subfragment GH4 contains a gene encoding a 1.1 kb mRNA. A corresponding cDNA ovD19C2 has been isolated from a *Drosophila* ovary cDNA library.

Northern analysis shows that this gene is expressed at low levels during most developmental stages, while the transcription is enhanced in the ovary and testis. High level mRNA of this gene is also present in the early embryos. In situ hybridization shows that the enhanced transcription occurs in nurse cells and spermatocyte cysts. The high level mRNA of this gene in the early embryo imply active synthesis of this message must occur in the nurse cells, and it is transported and stored in the oocyte.

Sequencing analysis shows this gene contains three introns. The longest open reading frame of the cDNA encodes a protein of 227 amino acid residues. The amino acid sequence of the predicted protein shows strong similarity to the human ubiquitin carboxyl terminal hydrolases. It has 51.1% identity to the human ubiquitin carboxyl terminal hydrolase isozyme L3 (UCH-L3), and 47.% identity to isozyme L1 of the same enzyme. The predicted protein of ovD19C2 also has many important features which are common in this enzyme family.

Attempts were made to mutate this gene by imprecise excision of the P-transposon from the strain D19. 45 lines which are homozygous lethal for chromosome 3 have been obtained. Further characterization of these lines will provide very important information about the gene cloned.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Mary Bownes for her help, encouragement and advice during this project. I also thank my advisor Dr. David Finnegan for his constructive suggestions.

I would like to thank all past and present members of Bownes' group for their help, care, and tolerance of the mess I have made: especially Alberto, Alan, Andrea, Brenda, Claire, Debbie, Elaine, Gregor, Katrin, Kathleen, Roger and Susan.

I thank the media and washing-up room workers, especially Ruby and Joan for not only always providing clean glassware but also making me known many thing about this society that I could have never learnt in the lab. I thank Graeme Brown and Frank Johnstone for their excellent work for processing all the photographs.

I thank Andrew Mitchison for his carefully correcting my somewhat terrible English to make it understandable.

I thank Roger and Brenda for their jokes, and serious discussion from Jocky to enhancer binding proteins. And thank Roger to let me know many things and places in Scotland.

I thank my friends Yan-Biao, Ru-Bing, Lin-Qi for their support and chat that make life a little bit easier.

I thank the British Council for financial support of this project.

Finally, from the bottom of my heart, I thank my parents and my sister for their love and support.

ABBREVIATIONS

| | |
|-------------------|--|
| amp | Ampicillin |
| APS | Ammonium persulphate |
| ATP | Adenosine-5'-triphosphate |
| BCIG(X-gal) | 5-bromo-4-chloro-e-indoyl-galactoside |
| bp | Base pair |
| BSA | Bovine serum albumin |
| °C | Degrees centigrade |
| cDNA | Complementary deoxyribonucleic acid |
| Ci | Curies |
| cm | Centimeter(s) |
| cpm | Counts per minute |
| dATP | Deoxyadenosine-5'-triphosphate |
| dCTP | Deoxycytosine-5'-triphosphate |
| dGTP | Deoxyguanosine-5'-triphosphate |
| dNTP | Deoxynucleotide-5'-triphosphate |
| dTTP | Deoxythymidine-5'-triphosphate |
| ddATP | 2'(3'-di) Deoxyadenosine-5'-triphosphate |
| ddCTP | 2'(3'-di) Deoxycytosine-5'-triphosphate |
| ddGTP | 2'(3'-di) Deoxyguanosine-5'-triphosphate |
| ddTTP | 2'(3'-di) Deoxythymidine-5'-triphosphate |
| dH ₂ O | Distilled water |
| DEAE | Diethylaminoethyl |
| DEPC | Diethyl pyrocarbonate |
| DMSO | Dimethylsulphoxide |
| DNA | Deoxyribonucleic acid |

| | |
|-------------------|--|
| DNase | Deoxyribonuclease |
| DTT | Dithiothreitol |
| EDTA | Diaminoethanetetra-acetic acid |
| EMS | Ethyl methyl sulphonate |
| fs | Female sterile |
| FSB | Formaldehyde sample buffer |
| g | Gram(s) |
| [³ H] | β-emitting isotope of hydrogen |
| HEPES | N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid |
| IPTG | Isopropyl-β-D-thiogalactopyranoside |
| k b | Kilobase |
| Klenow | Large fragment of DNA polymerase I |
| Krpm | Kilorevolutions per minute |
| λ | Lambda bacteriophage |
| L | Litre |
| M | Molar |
| m A | Milliampere(s) |
| mAP | Messenger Affinity Paper |
| mCi | Millicurie(s) |
| mg | Milligram(s) |
| ml | Millilitre(s) |
| mm | Millimetre(s) |
| mM/L | Millimoles per litre |
| mM | Millimolar |
| mmol | Millimole(s) |
| min | minute(s) |

| | |
|---------------------------|---|
| MOPS | Morpholinopropanesulphonic acid |
| mRNA | Messenger ribonucleic acid |
| NaI | Sodium iodine |
| NBT | 4-Nitro blue tetrazolium chloride |
| ng | Nanogram(s)) |
| nmol | Nanomole(s) |
| OLB | Oligo labelling buffer |
| OD | Optical density |
| [³² P] | β-emitting isotope of phosphorous |
| PEG | Polyethylene glycol |
| % | Percentage |
| pg | Picogram(s) |
| pH | -log ₁₀ (hydrogen ion concentration) |
| Poly (A) ⁺ RNA | Polyadenylated ribonucleic acid |
| psi | Pounds per square inch |
| RNA | Ribonucleic acid |
| RNase | Ribonuclease |
| rRNA | Ribosomal ribonucleic acid |
| rpm | Revolutions per minute |
| [³⁵ S] | β-emitting isotope of sulphur |
| SDS | Sodium dodecyl sulphate |
| TEMED | NNN'N'-tetra-methyl-1.2-diamino-ethane |
| Tris | Tris(hydroxymethyl)-amino-methane |
| Triton-X-100 | Octylphenoxypolyethoxyethanol |
| TCA | Trichloroacetic acid |
| U | Unit(s) |
| UTP | Uridine-5'-triphosphate |

| | |
|-----------------|--------------------------------------|
| UV | Ultraviolet |
| μCi | Microcurie(s) |
| μg | Microgram(s) |
| μl | Microlitres(s) |
| μM | Micromolar |
| μmol | Micromole(s) |
| V | Volts |
| v/v | Volume per volume |
| w/v | Weight per volume |
| X-phosphate | 5-Bromo-4-chloro-3-indolyl-phosphate |

CONTENTS

| | |
|-----------------|-----|
| Title | i |
| Dedication | ii |
| Declaration | iii |
| Abstract | iv |
| Acknowledgement | v |
| Abbreviations | vi |
| Contents | x |

CHAPTER ONE

| | |
|--|----|
| INTRODUCTION | 1 |
| Introduction | 2 |
| 1.1 Gonadogenesis | 3 |
| 1.2 Structure of Female Reproductive System of <i>Drosophila melanogaster</i> | 3 |
| 1.3 Oogenesis | 4 |
| 1.3.1 Differentiation of Germ Cells and Follicle Cells | 4 |
| 1.3.2 Growth of Egg Chambers | 6 |
| 1.4 Genetic and Molecular Studies of Oogenesis of <i>D. melanogaster</i> | 8 |
| 1.4.1 Female Sterile Mutations | 9 |
| 1.4.2 Maternal Effect Genes | 14 |
| 1.5 P-element Mediated Enhancer Detection | 20 |
| 1.5.1 P-element Mediated Gene Transfer | 20 |
| 1.5.2 Enhancer trapping | 21 |

CHAPTER TWO

| | |
|--|-----------|
| MATERIALS AND METHODS | 26 |
| 2.1 Solutions | 27 |
| 2.2 <i>Drosophila</i> Stocks and Strains | 29 |
| 2.3 <i>D. melanogaster</i> Life Cycle - Collection of Developmental stages | 30 |
| 2.3.1 Stages of <i>D. melanogaster</i> Development at 25°C | 30 |
| 2.3.2 Collection of Staged Eggs | 30 |
| 2.3.3 Collection of Early and Late 3rd Instar Larvae and Pupae | 30 |
| 2.3.4 Collection of Sexed Adults | 31 |
| 2.4.1 Microbial Strains | 31 |
| 2.4.2 Media | 31 |
| 2.5 Methods | 32 |
| 2.5.1 General Methods | 32 |
| 2.5.1.1 Phenol Extraction | 32 |
| 2.5.1.2 Precipitation of Nucleic Acids | 32 |
| 2.5.1.3 Restriction Endonuclease Digestion of DNA | 32 |
| 2.5.1.4 Extraction of DNA from Agarose Gels | 32 |
| 2.5.1.5 Ligation of DNA | 33 |
| 2.5.1.6 Preparation of Dialysis Tubing | 33 |
| 2.5.1.7 Transformation into <i>E. coli</i> | 33 |
| 2.5.2 Detection of β -galactosidase in | |

| | | |
|---------|--|----|
| | Adult Flies | 34 |
| 2.5.3 | Preparation of Genomic DNA from Adult Flies | 34 |
| 2.5.4 | Preparation of RNA from Different Developmental Stages and Tissues of <i>D. melanogaster</i> | 35 |
| 2.5.5 | Manipulation of Lambda | 35 |
| 2.5.5.1 | In Vitro Packaging of Phage Lambda DNA (Amersham) | 35 |
| 2.5.5.2 | Preparation of Lambda DNA | 35 |
| 2.5.5.3 | Plating of Lambda Libraries | 37 |
| 2.5.5.4 | Taking Plaque Lift | 37 |
| 2.5.6 | Preparation of Plasmid DNA | 37 |
| 2.5.6.1 | mini-scale | 37 |
| 2.5.6.2 | maxi-scale | 38 |
| 2.5.7 | Electrophoresis | 38 |
| 2.5.7.1 | Agarose Gel Electrophoresis of DNA | 38 |
| 2.5.7.2 | Agarose Gel Electrophoresis of RNA | 39 |
| 2.5.7.3 | SDS-Polyacrylamide Gel Electrophoresis of Proteins | 39 |
| 2.5.7.4 | Polyacrylamide Gel for DNA Sequencing | 40 |
| 2.5.8 | Transfer of Nucleic Acid to Membranes | 40 |
| 2.5.8.1 | Southern Blot | 40 |
| 2.5.8.2 | Northern Blot | 41 |
| 2.5.9 | Labelling Nucleic Acid | 41 |
| 2.5.9.1 | Radioactive labelling DNA by Random Priming | 41 |

| | | |
|----------|---|----|
| 2.5.9.2 | Digoxigenin Labelling DNA by Random Priming | 41 |
| 2.5.9.3 | Digoxigenin Labelling RNA by In Vitro Transcription | 41 |
| 2.5.10 | Hybridization of Labelled Probes to Membrane Filters | 42 |
| 2.5.11 | Signal Detection for Membrane Filters and Sequencing Gels | 42 |
| 2.5.11.1 | Autoradiography | 42 |
| 2.5.11.2 | Immuno-detection of Digoxigenin Labelled Probes | 42 |
| 2.5.12 | Generation of Unidirectional Deletions with Exonuclease III | 43 |
| 2.5.13 | DNA Sequencing by Dideoxynucleotide Chain Termination Method | 44 |
| 2.5.14 | <i>In Situ</i> Hybridization to Third Instar Polytene Chromosome Preparations | 46 |
| 2.5.14.1 | Preparation of Third Instar Polytene Chromosomes | 46 |
| 2.5.14.2 | Prehybridization of Polytene Chromosome Preparations | 47 |
| 2.5.14.3 | Hybridization | 47 |
| 2.5.14.4 | Washing Steps After Hybridization | 47 |
| 2.5.14.5 | Microscopy and Chromosome Staining | 48 |
| 2.5.15 | <i>In Situ</i> Hybridization for Localization of Specific RNA in <i>Drosophila</i> Ovaries, | |

| | | |
|----------|---------------------------|----|
| | Testes and Embryos | 48 |
| 2.5.15.1 | Tissue Preparation | 48 |
| 2.5.15.2 | Fixation | 48 |
| 2.5.15.3 | Pretreatment | 49 |
| 2.5.15.4 | Hybridization and Washing | 49 |
| 2.5.15.5 | Signal Detection | 50 |

CHAPTER THREE

| | | |
|-------|---|----|
| | IN SITU DETECTION OF ENHANCERS ACTING IN ADULT GONADS | 51 |
| 3.1 | General Description of the <i>LacZ</i> staining Patterns | 52 |
| 3.2 | LacZ Staining Patterns of the Ovary | 52 |
| 3.2.1 | Somatic Specific Staining Patterns | 52 |
| 3.2.2 | Germ-Line Specific Staining Patterns | 54 |
| 3.3 | LacZ Staining Patterns of Male Gonads | 55 |
| 3.4 | Comparison of the Expression Patterns β -galactosidase between Male and Female Gonads | 55 |
| 3.5 | The Expression Pattern of β -galactosidase in Strain D19 | 57 |
| 3.6 | Discussion | 58 |

CHAPTER FOUR

| | | |
|-----|---|----|
| | CLONING OF THE DNA SEQUENCE FLANKING THE P[LAC, RY+] INSERTION IN STRAIN D19 | 61 |
| 4.1 | Introduction | 62 |
| 4.2 | Construction and Screening of the D19 Genomic Library | 62 |

| | | |
|-----|-----------------------------------|----|
| 4.3 | Searching for Transcription Units | 64 |
| 4.4 | Discussion | 65 |

CHAPTER FIVE

| | | |
|---|---|-----------|
| ISOLATION AND SEQUENCING OF cDNAs | | |
| CORRESPONDING TO THE GENOMIC SEQUENCE FLANKING THE | | |
| P[LAC,RY+]A INSERTION | | 67 |
| 5.1 | Introduction | 68 |
| 5.2 | Screening of the Ovarian cDNA library | 68 |
| 5.3 | Analysis of the Homology of the Cloned | |
| | cDNAs to the Genomic Clone | 68 |
| 5.3.1 | Southern Analysis | 68 |
| 5.3.2 | <i>In Situ</i> Hybridization to the Third | |
| | Instar Larval polytene Chromosomes | 69 |
| 5.4 | Sequencing Analysis of the Genomic | |
| | Fragment GH4 and Corresponding | |
| | two cDNAs | 70 |
| 5.4.1 | Sequencing Strategy | 70 |
| 5.5 | Sequencing Analysis of OvD19C2 | 71 |
| 5.5.1 | Sequencing of OvD19C2 | 71 |
| 5.5.2 | Codon Usage of OvD19C2 | 72 |
| 5.5.3 | Comparison of OvD19 cDNA Sequence | |
| | to the Genomic Sequence | 72 |
| 5.5.4 | Computer Data Base Searching | 73 |
| 5.5.5 | Comparison of the OvD19C2 Protein | |
| | to the Human Proteins and Yeast | |
| | Proteins | 74 |
| 5.6 | Sequencing Analysis of OvD19C1 | 75 |

| | | |
|-----|------------|----|
| 5.7 | Discussion | 76 |
|-----|------------|----|

CHAPTER SIX

ANALYSIS OF THE EXPRESSION PATTERNS OF OVD19C2 AND OVD19C1B

| | | |
|--------|---|----|
| 6.1 | Introduction | 81 |
| 6.2 | Analysis of the Developmental Profile of OvD19C2 | 86 |
| 6.2.1 | Northern Analysis | 86 |
| 6.2.2 | <i>In Situ</i> Detection of the Transcript of OvD19C2 | 87 |
| 6.3 | Northern Analysis Using the OvD19C2 Probe | 88 |
| 6.3.1 | Comparison of the Transcripts Detected by the OvD19C1 and OvD19C2 Probes | 88 |
| 6.3.2. | Expression Pattern of OvD19C1 during Development | 88 |
| 6.4 | Discussion | 88 |

CHAPTER SEVEN

PRELIMINARY EXPERIMENTS TOWARDS

UNDERSTANDING OF THE FUNCTION OF OVD19C2

| | | |
|-------|--|----|
| 7.1 | The Expression of OvD19C2 in <i>E. coli</i> | 93 |
| 7.2 | Excision of the P[<i>lac</i> , <i>ry</i> ⁺]A insertion from D19 | 93 |
| 7.2.1 | Mobilizing P[<i>lac</i> , <i>ry</i> ⁺]A transposon by Crossing to P[<i>ry</i> ⁺ , Δ 2-3](99B) | 93 |
| 7.2.2 | Did Deletion Occur? Analysis by Southern Blotting | 95 |

| | | |
|-------------------------|------------|-----|
| 7.3 | Discussion | 96 |
| FINAL DISCUSSION | | 98 |
| REFERENCES | | 104 |

CAHAPTER ONE

INTRODUCTION

Introduction

Oogenesis is a very complex process. During this process the germ line cells differentiate to produce the female gamete, namely the oocyte. To produce the oocyte requires cell determination, cell differentiation and cell interaction; both between germ-line cells, and between germ-line cells and somatic cells. During oogenesis large amounts of maternal information are stored in the oocyte, these are vital for future embryonic development. Nutritional and structural materials are also produced during oogenesis and stored in the oocyte, and are used to produce mature eggs (e.g. yolk proteins and chorion proteins). The fruit fly, *Drosophila melanogaster*, due to its well characterised genetics, is one of the best systems for the study of oogenesis.

1.1 Gonadogenesis

The gonads of both sexes of *D. melanogaster* are composed of germ-line cells, which are descendants of the pole cells in the embryo, and mesoderm-derived gonadal somatic tissues.

The pole cells are the first obvious differentiated cells in the embryo. After the 8th nuclear division a group of nuclei migrate to the posterior end of the embryo beneath the plasma membrane, where they are surrounded by the pole plasm containing the polar granules; believed to contain the determinants for pole cell formation (this will be discussed later in 1.3.2). After cellularization, about ten pole cells are formed. Following further mitotic divisions, the number of pole cells reaches 40 to 60. During gastrulation these pole cells are passively carried into the posterior pocket of the midgut primordium as they migrate along dorsally towards to anterior of the embryo. At about 7 hours post fertilisation, about 50-60% of these pole cells penetrate through the midgut primordium. They move in to the body cavity and gain their final positions dorsal-posterior to the midgut primordium, where they are then enclosed by a monolayer of small mesodermal cells; which later differentiate to be the somatic gonadal tissues (Sonnenblick, 1941).

The difference between male and female gonads marks the earliest sign of sexual dimorphism. This difference is characterised by the numbers of germ cells included in the gonads. The male gonad usually contains 5 to 7 germ cells whereas the female gonad contains 9 to 13 germ cells.

Little is known about the molecular and cellular events involved in the establishment of the embryonic gonads, except those maternally expressed genes whose products are required for the formation of pole cells and development of abdomen. Mutations at these loci usually disrupt pole cell formation and hence cause the "grandchildless" phenotype.

1.2 Structures of the Female Reproductive System of *Drosophila melanogaster*

In adult females of *D. melanogaster*, a pair of ovaries is located in the abdomen, which consists of a clusters of parallel ovarioles, usually 15-20. The ovaries are surrounded by the peritoneal sheath with a network of muscles. Eggs maturing in the ovarioles are pushed through the common oviduct into the uterus, where they are fertilized by sperm stored in the receptacle and spermathecae. After fertilization, eggs are laid through the vulva (King, 1970)(Fig.1.1)

An ovariole can be divided into two regions: these are called the germarium and the vitellarium. At the tip of each ovariole is the germarium, and the remainder of the ovariole is occupied by the vitellarium. A series of egg chambers at different stages of development can found in an ovariole. King *et al.* (1956) divided the oogenesis of *Drosophila* into 14 stages (Figure 1.2). In the germarium, both germ cells and somatic cells differentiate and at the posterior end of the germarium, the stage 1 egg chamber is formed, consisting of one oocyte and 15 nurse cells surrounded by a monolayer of about 80 follicle cells. From stage 2 to stage 7 in the vitellarium, the shape of the egg chamber changes from spherical to oval. Because there is no yolk accumulation in the oocyte during these stages, they are called previtellogenic. From stage 8 onwards, oocyte growth occurs due to the uptake of yolk and other materials, and these stages are classified as vitellogenic stages. The nurse cells eventually degenerate, and a mature egg can be found in a stage 14 egg chamber. At this time both the anterior-posterior and dorsal-ventral polarities of the egg have been established.

1.3 Oogenesis

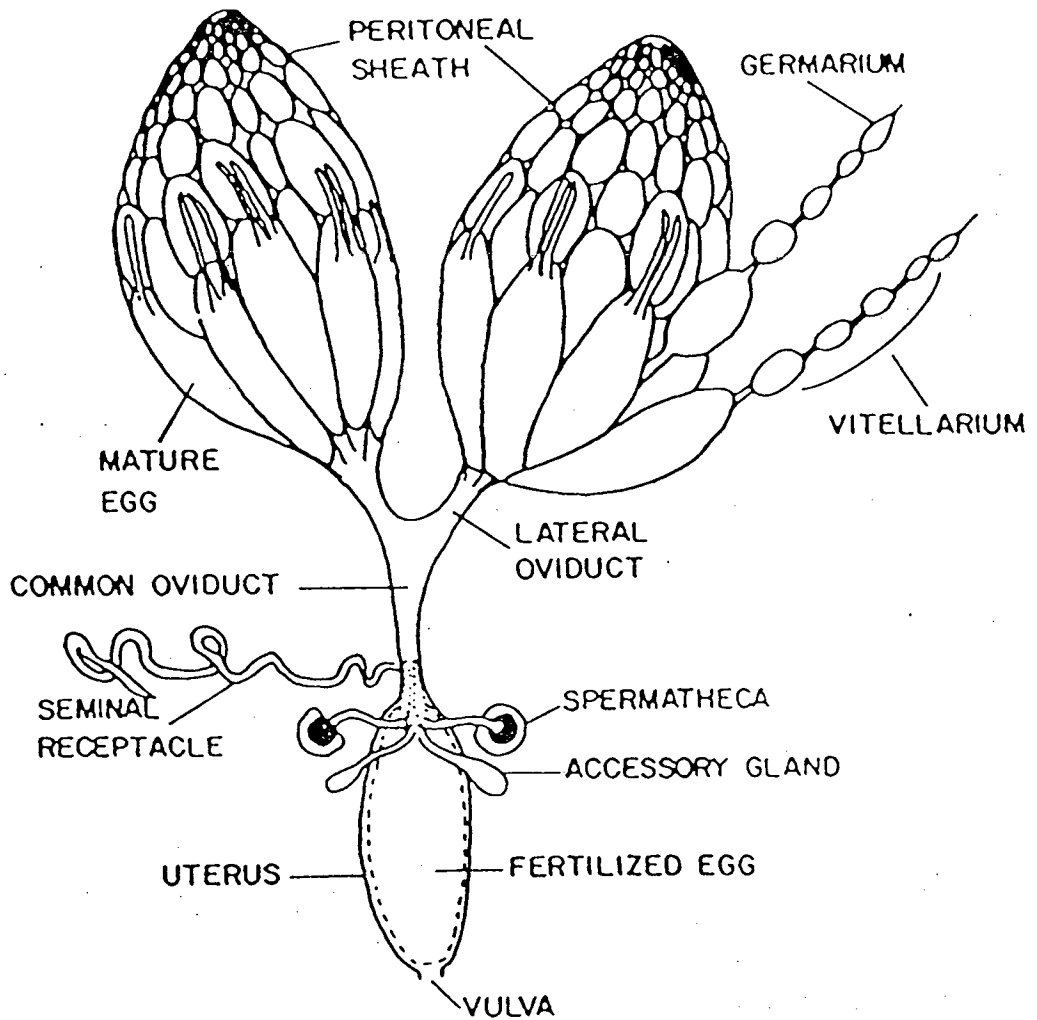
1.3.1 Differentiation of Germ Cells and Follicle Cells

Based on electron microscopic observations, the germarium can be divided into three regions: Region 1, Region 2 and Region 3 (Figure 1.3) (Mahowald and Strassheim, 1970). Region 1 is a mitotically active area situated at the tip of the germarium, where individual stem cells, cystoblasts and clusters of cystocytes can be observed. In Region 2, the cystocyte clusters are separated by the invasions of mesodermal derived prefollicle cells. In Region 3, the

Figure 1.1

Structure of the Female Adult Reproductive System of *Drosophila melanogaster* (from Bownes, 1982).

FIGURE 1.1



prefollicle cells finally form a monolayer of cells which surrounds the complex of 15 nurse cells and the oocyte derived from a single stem cell.

Germ cell differentiation in *D. melanogaster* begins after pupariation. From genetic mosaic analysis, it has been established that there are at least two stem cells per ovariole (Weischaus and Szabad, 1979). After each division of the stem cell, one daughter cell remains as the stem cell, while the other differentiates as a cystoblast. The mechanism controlling this differentiation remains unknown. The stem cell grows after each division, until its original size is regained, while the cystoblasts divide continuously without increasing their size, until the 16 cell cystocyte cluster is formed. King (1970) postulated that the determination of the fate of the two daughter cells involves their positions in the germarium. Attachment to the terminal filament might have some effect on maintenance of the status of the stem cell, because some secretory structures were observed in the terminal filament.

Another important event in germ cell differentiation is the determinative decision between the oocyte and the nurse cells. The cystoblast divides following a 2^n rule for four divisions, which results in a 16 cell cystocyte cluster. Due to incomplete cytokinesis, cells within a cystocyte cluster are connected to each other by intercellular bridges. Among the 16 cystocytes two cells have 4 bridges, two cells have three bridges, four cells have two bridges and eight cells have one bridge (Brown and King, 1964). Although all 16 cells undergo synchronous DNA synthesis, synaptonemal complexes are only observed in the two cells which have four intercellular bridges, and the one of them which finishes meiosis becomes the oocyte (Koch *et al.*, 1967; Koch and King, 1969; Day and Gell, 1976). The mechanism controlling their differentiation remains to be determined.

The prefollicle cells cannot be distinguished from other somatic cells in the female gonad until the pupal stage. In the germarium of the adult fly, two types of prefollicle cells can be observed. In Region 1 and Region 2, squamous prefollicle cells surround the stem cells and the cystocytes. The arrangement of the

Figure 1.2

Structure of the Ovariole, Showing the Stages of Oogenesis in *Drosophila melanogaster* (based on King, 1970)

Numbers are stages of oogenesis

DA = dorsal appendage

G = germarium

F = follicle cell

N = nurse cell

O = oocyte

FIGURE 1.2

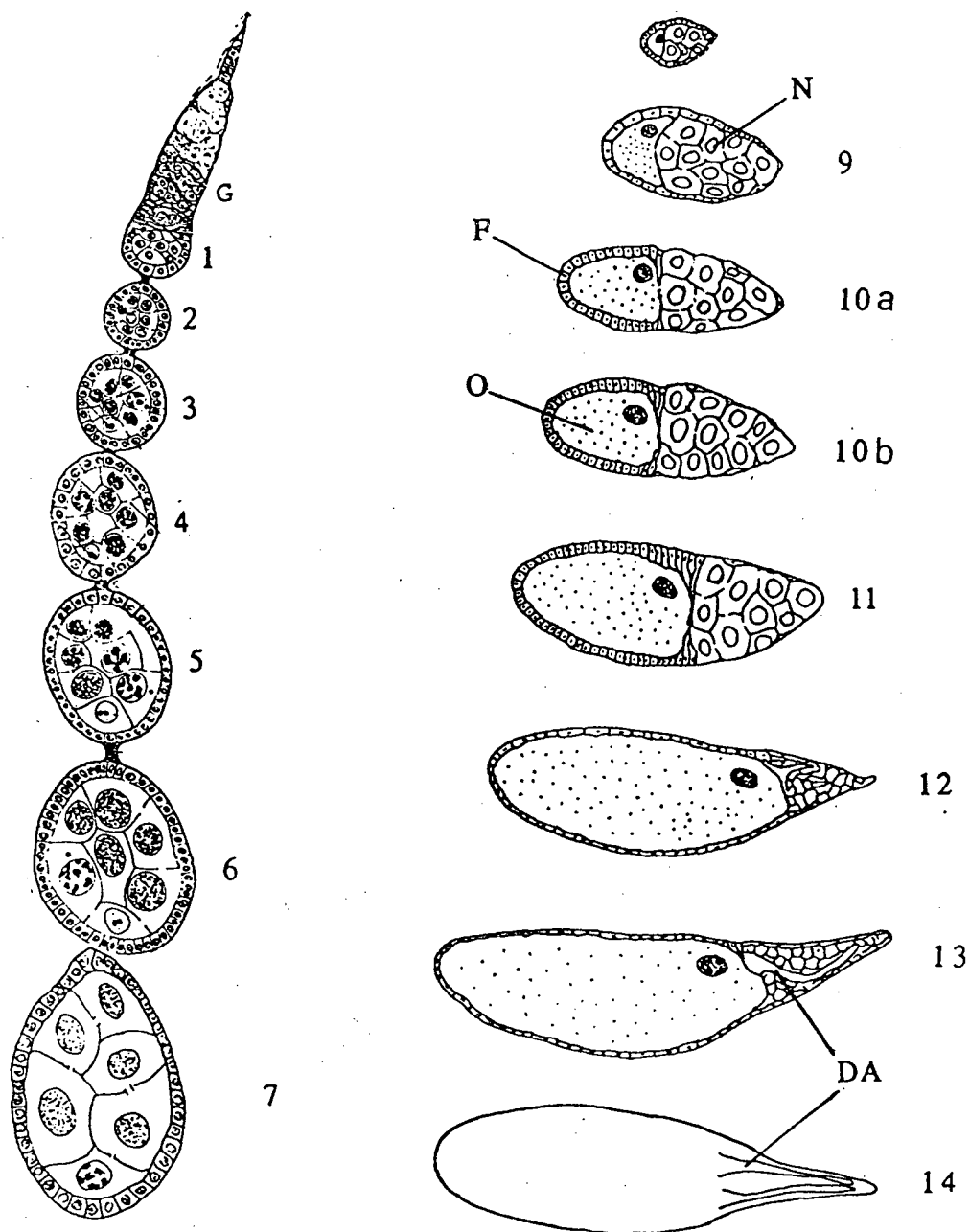
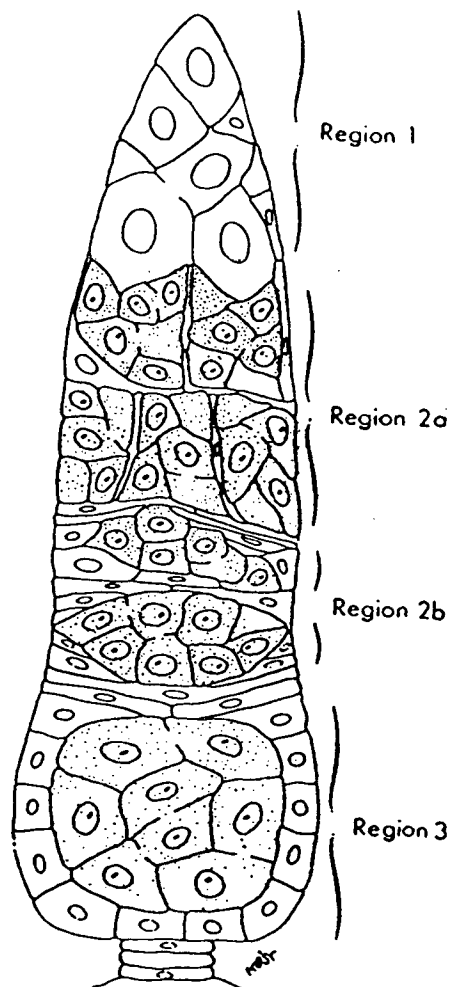


Figure 1.3

Structure of Germarium (from Mahowald and Strassheim, 1970)

FIGURE 1.3



prefollicle cells is irregular at this stage, and a thin layer of one or more cells can be found which separates the cystocyte clusters. The cell layer decreases while the cysts are moving towards the posterior along the ovariole. A second group of prefollicle cells are found in Region 3. They are the 80 cells which form the monolayer surrounding the spherical cyst (King, 1970). How the prefollicle cells in Region 1 and Region 2 change to a monolayer in Region 3 is not clear. It is likely that when the cyst grows, it could stretch the prefollicle cells surrounding it (Mahowald and Strassheim, 1970).

1.3.2 Growth of the Egg Chambers

As the egg chambers enter the vitellarium, the nurse cells start DNA synthesis. As a result of endomitosis the nurse cells become polyploid. Although DNA synthesis is not synchronous by the end of stage 5, all nurse cells have a ploidy of 64C. The extent of polyploidization between nurse cells starts to vary after stage 5. By the end of stage 9, the ploidy can reach 1024C in some nurse cells (King, 1970). Such an intensified polyploidization implies gene amplification, and is very much needed for the nurse cells to provide maternal materials for the oocyte. Almost all RNA in the mature egg is synthesised in the nurse cells (King and Burnet, 1959; Zalokar, 1965; Mahowald and Tiefert, 1970). Based on autoradiographic studies (King and Burnett, 1959), it was showed that RNA synthesis occurred in all nurse cells within an egg chamber of certain stages, and most of the nuclear RNA later entered the cytoplasm was then transported into the oocyte through the intercellular bridges (Zalokar, 1960).

The oocyte increases in volume as a result of the accumulation of materials provided by nurse cells and yolk proteins taken up from haemolymph and follicle cells, and is itself transcriptionally inactive for much of oogenesis. A brief period of transcription is detected in the oocyte between stage 9 and 10 (Mahowald and Tiefert, 1970), and so far only one gene is known to be expressed in the oocyte nucleus. It is called *fs(1)K10*, and is involved in the formation of the dorsal-ventral polarity of the follicle cells around the anterior end of the oocyte (Prost et al., 1988). These follicle

cells will specify the structure of the egg shell in the anterior region. This gene will be discussed later in 1.4.1 in more detail.

During the process of vitellogenesis, a large amount of nutritional and structural material is accumulated in the oocyte. The yolk proteins (YPS) are the main components of these materials. In *D. melanogaster* there are three of yolk proteins: namely YP1, YP2 and YP3, which are synthesized in fat-bodies and follicle cells (Kambysellis, 1977; Bownes and Hames, 1978). The synthesis of yolk proteins is induced by 20-hydroxyecdysone, while the Juvenile Hormone regulates the uptake of yolk proteins from the haemolymph into the oocyte (Kambysellis, 1977; Postlethwait *et al.*, 1976).

During vitellogenesis, follicle cells also secrete the vitelline membrane and chorion proteins to form the coat of the egg. From stage 1 to stage 6, follicle cells are mainly involved in their own proliferation. Mitoses end at stage 6 of oogenesis, when approximately 1000-1200 follicle cells surround each egg chamber. About 90% of follicle cells migrate posteriorly at stage 8-9, to form a high columnar monolayer around the oocyte, whilst most of the remaining 10% follicle cells surrounding nurse cells stretch into a thin, squamous epithelium. At about stage 9, a group of 6-10 follicle cells located at the most anterior end of the egg chamber, migrate between nurse cells to antero-dorsal end of the oocyte, where they differentiate into the border cells, which are responsible for the formation of micropyle. This is the place where the sperm will enter the egg. By the end of stage 10, a set of columnar follicle cells migrates and intercalates between nurse cells and the anterior end of the oocyte. Of these cells, a group of them secrete the two dorsal appendages, while the other group form the operculum where the larva will hatch. At the posterior end of the egg chamber, the follicle cells form the aeropyle, a structure thought to be required for the respiration of the embryo (Margaritis *et al.*, 1980; Margaritis, 1985)

At the posterior end of the egg, a type of special material can be found. This is composed of polar granules. They first appear at about stage 9 to stage 10 of oogenesis. During stage 12 and 13, they

attach to each other to form a long chain. Based on electron microscopic observations, it is found that these polar granules are composed of small, fibrous bodies (Mahowald, 1962, 1971a). The polar granules consists of RNA and protein (Mahowald, 1971b), and are believed to be involved in germ-cell determination. Several lines of evidence suggest that the polar granules contain the germ-cell determinants. Firstly, the pole cells which are normally formed at the posterior end of the embryo can be induced at the anterior end by injecting the pole plasm containing polar granules (Illmensee and Mahowald, 1976). Secondly, when the posterior end of the egg is irradiated with UV light, the formation of the pole cells is disrupted. This disruption can be rescued by injection of mRNA extracted from the posterior end of normal eggs (Okada and Koayaski, 1987). Finally, several mutations have been identified that disrupt biogenesis of polar granules, and no pole cells are formed in the embryos produced by these mutant mothers (Lehmann and Nüsslein-Volhard, 1986; Nüsslein-Volhard *et al.*, 1987)

1.4 Genetic and Molecular Studies of Oogenesis in *D. melanogaster*

Both genetic and molecular studies indicate that a large number of genes are involved during oogenesis to ensure normal ovarian development and embryogenesis. Based on genetic mosaic studies, at least 75% of all genes that are essential for the viability of the fly are also necessary for oogenesis (Garcia-Bellido and Robbins, 1983; Perrimon *et al.*, 1984). Hough-Evan *et al.*, (1980) analyzed the complexity of RNA species in *Drosophila* eggs, and concluded that 10% of all single copy sequences in the genome are found in ovarian RNA.

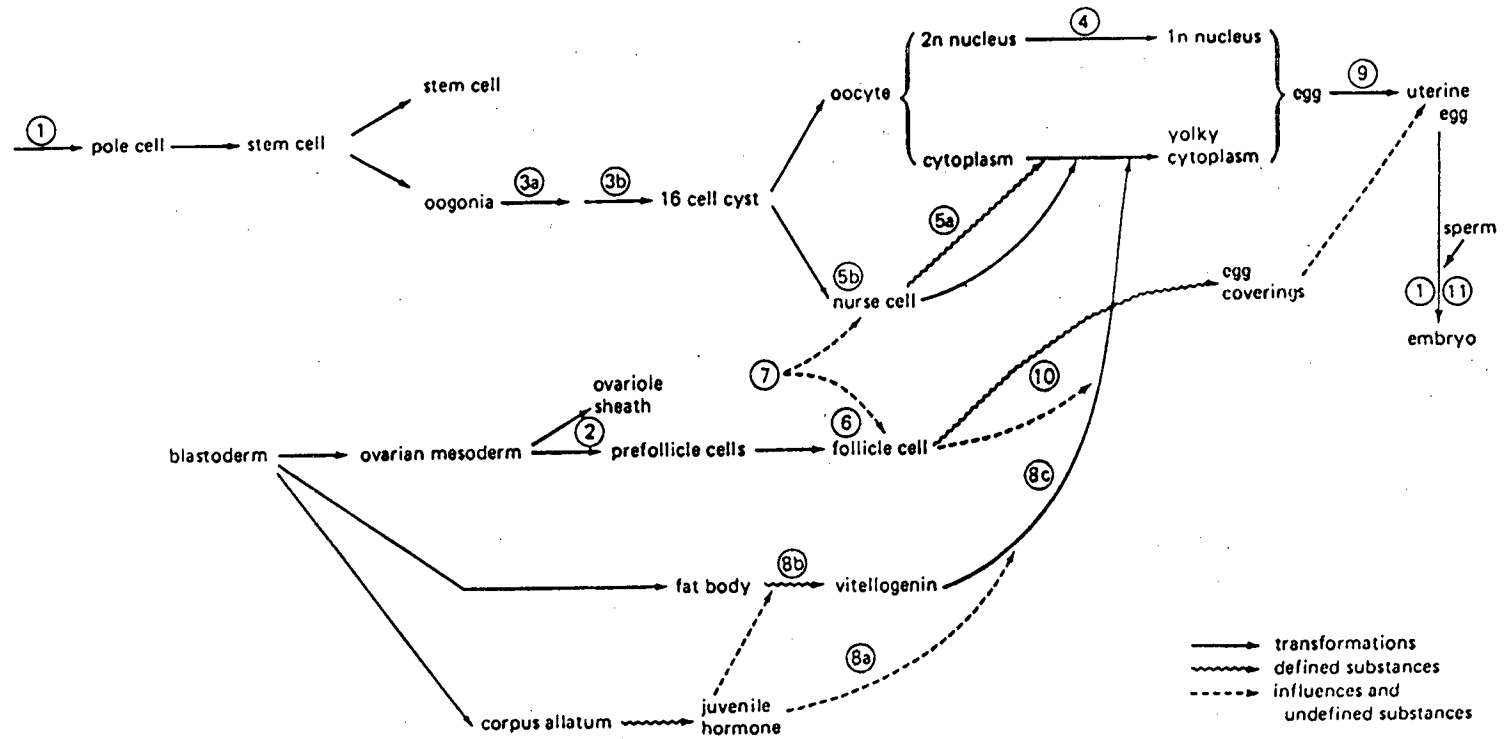
In the past two decades, great attempts have been made to screen the *Drosophila* genome by means of X-ray and EMS mutagenesis to identify female sterile mutations (fs) (Bekken, 1973; Gans *et al.*, 1975; Mohler, 1977; Komitopoulou *et al.*, 1983; Schüpbach and Weischaus, 1989). As a result, a large number of loci have been identified which, when mutated, either cause abnormal

Figure 1.4

A diagram illustrating the mode of origin of the *Drosophila* oocyte and its interaction with other cell types during its development (from King, 1975).

Circled numbers refer to each stage which can be affected by various mutations.

FIGURE 1.4



ovarian development, or have a maternal effect on embryogenesis. These fs mutations can be divided into two classes according to the defects they cause. Class A comprises those mutations which disrupt normal ovarian development, or produce morphologically defective eggs. Class B are called maternal effect mutations, which although they produce morphologically normal eggs, fail to provide correct or necessary information for embryonic development. Figure 1.4 is a summary of the oogenesis of *D. melanogaster* and the involvement of various genes in each step of this process.

Table 1.1 is a list of some fs mutations in Class A. As shown in this table, oogenesis can be disrupted at different developmental stages and these defects can be observed in different cell types. Several examples of the Class A mutation are discussed below.

1.4.1 Female Sterile Mutations

A. Ovarian Tumours

Several genes (*ovo*, *sans fille* (*snf*), *Sex-lethal* (*Sxl*), and *ovarian tumour* (*otu*)) required for female cystocyte cyst formation have been identified, which when mutated, and in certain allelic conditions, produce "tumorous egg chambers", i.e. egg chambers are full of irregular proliferating cells and abnormal cysts (Pauli and Mahowald, 1990). A common feature of these mutations is that germ cells in these tumorous ovaries all show morphological similarity to the spermatocyte cysts, indicating that these genes must be required for sex-determination in germ-line cells. Evidence from genetic and molecular studies shows that in the tumorous ovaries caused by *Sxl*, *snf* and *otu* mutations, the gene *Stellate*, normally expressed in male germ-line cells, is found to be expressed in female germ-line cells in these mutants (Steinmann *et al.*, 1989), suggesting that these germ cells are undergoing male differentiation.

The germ-line cells of both sexes require the correct sexual identities of surrounding gonadal somatic tissues to ensure normal development. An incompatible combination of germ-line cells and soma usually causes cell death in early stages, or ovarian tumour development in late stages of development. The male and female

Table 1.1

Classification Of Female Sterile Mutants
Based on King and Mohler (1975)

Category I Mutants: Female Sterile Mutants With Ovarian Defects

| <u>Class</u> | <u>Mutant</u> | <u>Defect</u> | <u>Reference</u> |
|---------------------------------|----------------------------------|---|--|
| 1. Germ cell determination | <i>grandchildless</i> | Mutants lay eggs that develop into adults generally lacking gonads | Thierry-Mieg <i>et al.</i> (1972) |
| | <i>Bicaudal-D</i> | Role in differentiation of oocyte from nurse cells. Involved in nanos localisation | Mohler and Weischaus (1986) Wharton and Struhl (1990) |
| 2. Differentiation of ovarioles | <i>fs(2)A10</i> | disruption of ovariole differentiation | Bekken (1973) |
| 3. Cystocyte divisions | <i>ovo</i> | normal egg-chamber development | Mével-Ninio <i>et al.</i> , (1989) |
| | <i>ovarian tumour sans fille</i> | is disrupted, cystocyte cysts composed of small, un-differentiated cells resembling spermatocytes | Stoto and King (1988) Steinmann-Zwicky (1988) |
| 4. Meiosis | <i>$\alpha(3)G^8$</i> | defective in chromatid crossing over | Hall (1972) |

| | | | |
|------------------------------|------------------|---|---|
| 5. Nurse cell | <i>chickadee</i> | mutant females have undersized oocytes that remain attached to persistent nurse cells | Cooly <i>et al.</i> (1990) |
| | <i>kelch</i> | | |
| | <i>bobbed</i> | partial deletion in genes coding for rRNA. Eggs are produced at a reduced rate by females homozygous for <i>bb</i> . Pleiotropic mutation affecting all cells. | Mohan and Ritossa (1970) |
| 6. Follicle cell development | <i>tiny</i> | Follicle cell migration affected and consequently their synthetic activity. Homozygous flies are late to eclose and have a small body covered in thin, short bristles | King (1970) |
| 7. Vitellogenesis | <i>fs(1)1163</i> | homozygous females are defective in yolk Protein-1 production, though this might not be the primary | Gans <i>et al.</i> , (1975) Sauders and Bownes, (1986) |

| | | | |
|-------------------------|-------------------|---|--|
| 8. Ovulation | <i>hairy wing</i> | defect. failure of mature oocytes to pass from the oviduct to the uterus. Has other pleiotropic effects. | Holzworth <i>et al.</i> (1974) |
| 9. Chorion synthesis | <i>ocelliless</i> | mature oocytes produced which contain greatly reduced amounts of several major chorion proteins including c36 and c38. Fewer eggs are laid. Pleiotropic mutation. | Spradling <i>et al.</i> (1979) |
| 10. Egg shape | <i>fs(1)K10</i> | dorsalisation of egg (and embryo). Ventrally fused eggs with enlarged chorionic appendages. | Wieschaus <i>et al.</i> , (1978) |
| | <i>gurken</i> | ventralisation of egg and embryo | Haenlin <i>et al.</i> , (1987) |
| | <i>torpedo</i> | | Schüpbach and Wieschaus (1989) Price <i>et al.</i> , (1989) |

germ-line cells appear to take different sex determination pathways. From the transplantation experiments, it has been found that the 1X:2A germ-line cells undergo male differentiation, irrespective of whether they are transplanted into male or female somatic environments. However, spermatogenesis is arrested at the spermatocyte stage if the surrounding somatic tissue is 2X:2A. This implies that sex-determination in 1X:2A germ-line cells is cell-autonomous, whereas the correct sexual type of surrounding somatic cells is still required to maintain appropriate differentiation. In contrast, sex-determination in the 2X:2A germ-line cells is non-autonomous; they choose their sexual fate according to the sexuality of the surrounding soma. Then if the 2X:2A germ-line cells are implanted into 1X:2A somatic surroundings, they undergo male differentiation, but it ceases at the spermatocyte stage. When 2X:2A germ-line cells homozygous for a loss-of-function allele of *Sxl* are transplanted in a 2X:2A embryo, the germ cells undergo spermatogenesis. In contrast, in the presence of the gain-of-function allele *Sxl*^{M#1}, the 2X:2A germ-line cells enter oogenesis, even when they are transplanted into an 1X:2A embryo (Steinmann-Zwicky *et al.*, 1989). These facts suggest that somatic influence on sex-determination of germ-line cells is passed through *Sxl*, and the genes described above must be under the control of *Sxl*. Combining the information described above, the ovarian tumour phenotype caused by these mutations, must result from sex-transformation in female germ-line cells, which then fail to develop normally in the incompatible somatic environment.

B. The Female Sterile Mutation, *fs(1)K10 (K10)*

The female sterile mutation *fs(1)K10* was originally described as affecting egg-shell structure. Females homologous for this mutation produce eggs with a dorsalized egg shell, and most of these eggs cannot be fertilized. A small proportion (1-2%) of fertilized eggs produce dorsalized embryos (Wieschaus *et al.*, 1978). Abnormal development can be traced to oogenesis in the follicle cells. The wildtype oocyte nucleus is acentric and lies near the surface at the antero-dorsal side of the oocyte. The overlaying

follicle cells on the side near the oocyte nucleus are much thicker than those on the opposite side. In *K10* egg chambers, all the follicle cells surrounding the anterior side of the oocyte become thickened in the same way as the follicle cells (normally found on the dorsal side) in spite of the fact that the oocyte nucleus still retains its acentric position. Although the phenotype of the *K10* mutation only appears in the follicle cells of somatic origin, it has been demonstrated by pole cell transplantation and somatic recombination, that the *K10* mutation is solely germ-line dependent (Wieschaus *et al.*, 1978). Molecular analysis has revealed that the *K10* gene is transcribed and its product translocated in the oocyte nucleus during oogenesis (Haenlin *et al.*, 1987, Prost *et al.*, 1988). The morphological change in the follicle cells caused by the *K10* mutation suggest that the *K10* gene must provide a morphogenic cue which, either directly or indirectly, influences the differentiation of follicle cells. Considering the acentric position of the oocyte nucleus and the asymmetric structure of the follicle cells, the morphogenic cue could possibly reach the follicle cells by random diffusion.

C. Chorion Genes

The genes encoding the chorion components that have been identified so far are clustered in two chromosomal regions: an X-linked gene cluster consisting of six genes (tandemly aligned at 7F1-2 on the X-chromosome), and another four genes located at 66D11-15. Ultrastructural and biochemical studies have revealed that the egg-shell is composed of multiple layers, and each layer is made of different proteins encoded by the genes mentioned above (Margaritis, 1985). The secretion of chorion protein takes a fairly short time during oogenesis (stage 10-14) (Griffin-Shea *et al.*, 1982). The DNA sequences of the chorion gene clusters are preferentially amplified to 15- or 60-fold of the haploid copy, before these genes are transcriptionally active (Griffin-Shea *et al.*, 1982). Such an extent of amplification appears to be a mechanism allowing a single copy gene to produce large amounts of short-lived transcripts in a very short period. The structures are different at

different regions of the egg-shell; those required by the egg for respiration, fertilization, larva hatching, etc. The complexity of the egg shell requires these genes, whose products are structural materials of the egg shell, to be precisely regulated both temporally and spatially.

The X-linked chorion genes encode six transcripts including two major chorion proteins (s36 and s38) and four putative chorion proteins (Parks *et al.*, 1986). The time (stage 10-13) of transcription of the X-linked chorion genes corresponds well with the formation of the innermost chorionic layer (ICL) and the endochorion (Spradling and Mahowald, 1979; Griffin-Shea *et al.*, 1982; Parks *et al.*, 1986). Mutations disrupting one or several X-linked chorion genes produce egg-shell defects primarily in the ICL and endochorion (Digan *et al.*, 1979). The transcription of the four chorion genes on chromosome 3 cannot be detected until stage 13. One gene encoding the s15 protein appears to have a biphasic transcription; its mRNA can be detected at stage 8-9, whereas protein synthesis does not take place until stage 13. The premature transcript is degraded before translation starts. The time of expression of these genes suggests that their products are constituents of the outer layer of the endochorion and exochorion.

In situ hybridization experiments have provided further details of the temporal and spatial expression pattern of these genes (Parks and Spradling, 1987). The time sequence of the expression of these chorion genes is consistent with the previous description. In addition, the results from in situ hybridization are very helpful in understanding the spatial regulation of the chorion genes. Of the ten transcripts from the two chorion gene clusters studied, six abundant chorion protein mRNAs are transcribed in most of the follicle cells. This more or less uniform transcription may reflect the general demand of these major chorion proteins, and is consistent with the location of their presence. On the other hand, three transcripts encoding minor chorion proteins are found transcribed in spatially restricted patterns, which are believed to specialize particular structures of the egg-shell, such as the dorsal appendages, operculum, micropyle and posterior aeropyle.

D. Yolk Protein Genes

Three yolk protein genes, (*yp1*, *yp2* and *yp3*) have been cloned. The genes *yp1* and *yp2* are located at 8F-9A, and *yp3* is at 12BC on the X-chromosome (Postlethwait and Jowett, 1980; Barnett *et al.*, 1980). The three yolk proteins are expressed in fat body and ovarian follicle cells of adult females, and are taken up by the oocyte as storage proteins at about stage 8-10 during oogenesis (Bownes and Hames, 1978; Brennan *et al.*, 1982). The *yp1* and *yp2* genes have been characterized in more detail than the *yp3* gene in terms of their regulation. The two genes (*yp1*, *yp2*) are transcribed divergently from each other, and a 1224 bp DNA sequence separates the two genes from their cap site (Hung and Wensink, 1983). By germ-line transformation, two enhancers have been identified. One is necessary to confer the fat-body-specific expression of *yp1* and *yp2*, while the other is required for follicle-cell-specific transcription (Garabedian *et al.*, 1985, 1986).

Expression of *yp* genes is strictly sex-specific and is under the control of the somatic sex-determination hierarchy. In the case of fat-body expression, continuous activity of the gene *tra-2* is required, as demonstrated by temperature-shift experiments with female flies homozygous for the temperature-sensitive allele, *tra-2^{ts}*. At the permissive temperature, these flies develop as normal females, whilst at the restrictive temperature they will develop as pseudomales (Belote and Baker, 1982, 1983). The expression of *yp1* and *yp2* genes can be switched on or off, by shifting the *tra-2^{ts}* females between the permissive and restrictive temperatures (Belote *et al.*, 1985, Bownes *et al.*, 1987). In vitro experiments have shown that the *double-sex* (*dsx*) protein DSX binds specifically to the fat-body enhancer of *yp1* and *yp2* (Burtis *et al.*, 1991).

In the case of follicle-cell-specific expression, an element located between -159 and -340 of *yp2* cap site is found necessary and sufficient to initiate the transcription of both the *yp1* and *yp2* genes (Garabedian *et al.*, 1985). A second regulatory element lies within the first exon of *yp2*, which acts across the *yp2* promoter to intensify *yp1* expression.

1.4.2 Maternal Effect Genes

In this class, mutant females produce morphologically normal eggs. However, embryonic development is disrupted in these eggs. In genetic screens for female-sterile mutations, a large number of maternal mutations have been recovered, and these show various defective phenotypes at different stages of embryogenesis. Some of these mutations are listed in table 1.2.

The most extensively studied maternal effect genes are those which determine embryonic pattern formation: the antero-posterior and dorsal-ventral body plans. There are four groups which determine the embryonic polarities and coordinate the body plan in the early embryo: the anterior group, the posterior group, the terminal group and the dorsal-ventral group. The first three groups determine the antero-posterior axis, while the last group defines the dorsal-ventral axis. The fact that mutations of different genes within a certain group produce identical or similar phenotype(s) suggests that these genes only specify pattern formation of the region where the mutations show their defects.

A. The Anterior Group

All the mutations in the anterior group produce embryos with some anterior structures (e.g. head and thorax) deleted (Nüsslein-Volhard, C. 1991). In this gene group, the *bicoid* (*bcd*) gene has been well characterized. Weak alleles cause the head-fold to be shifted posteriorly, while the strong alleles completely delete the head and thorax (Frohnhofer and Nüsslein-Volhard, 1986). It is believed that the product of the *bcd* gene functions as a morphogen which coordinates the anterior pattern formation. Evidence comes from experiments with cytoplasmic transplantation. The *bcd* phenotype can be rescued by injecting cytoplasm from the anterior end of freshly-laid wildtype eggs into the anterior end of mutant embryos. In addition, an ectopic head structure can be induced at the position where the *bcd* mRNA is injected (Frohnhofer and Nüsslein-Volhard, 1986). Molecular studies show that the *bcd* mRNA is transcribed in nurse cells during oogenesis, and is deposited in the cortex at anterior of the oocyte. In freshly-laid

Table 1.2

Classification Of Maternal Effect Mutations Which Disrupt Embryonic Development (Based On Schüpbach and Wieschaus, 1989)

Category B Mutants: Eggs Are Laid Which Fail To Develop Properly

| <u>Class</u> | <u>Mutant</u> | <u>Defect</u> | <u>Reference</u> |
|--|---------------|---|--|
| 1. Defects occurring during pre-syncitial stages | <i>fruh</i> | eggs show no visible signs of development; irregular cellular blastoderm | Schüpbach and Wieschaus (1989) |
| 2. Syncitial blastoderm arrest | <i>cortex</i> | narrow halo of clear cytoplasm appears at egg periphery | Schüpbach and Wieschaus (1989) |
| 3. Irregular cellularisation at blastoderm stage | <i>valois</i> | some embryos form pieces of cuticle 80-90% of embryos fail to cellularise | Schüpbach and Wieschaus (1989) |
| 4. Abnormal cell behaviour at onset of gastrulation - <u>Maternal Effect Loci Affecting Embryonic Pattern</u> | | | |
| A. Mutants Affecting the Formation of the Anterior-Posterior Pattern | | | |
| a) anterior group | <i>bicoid</i> | embryos lack anterior structures Primary anterior determinant) | Frohnhofer and Nüsslein-Volhard (1986) |

| | | | |
|--------------------|--------------------------------|--|---|
| | | | Driever and Nüsslein- Volhard (1988a, 1988b, 1989). |
| | <i>exuperentia swallow</i> | required for the correct localization of bicoid | Berleth <i>et al.</i> , (1988) |
| | | | Frohnhofer and Nüsslein- Volhard (1987) |
| | <i>hunchback</i> | maternal gap gene regulated by <i>bcd</i> Also expressed zygotically | Lehmann and Nüsslein- Volhard (1984); Tautz (1988); Tautz and Pfeifle (1989) |
| b) posterior group | <i>oskar</i> | deletion of abdomen | in: |
| | <i>pumilio</i> | (excluding telson) | Nüsslein- Volhard (1991) |
| | <i>nanos</i> | All except <i>nanos</i> and | |
| | <i>valois</i> | <i>pumilio</i> lack polar granules. | (<i>vasa</i> : |
| | <i>staufen</i> | (<i>nanos</i> may be the primary | In: |
| | <i>vasa</i> | posterior determinant) | Lasko and Ashburner (1990) |
| | <i>bicaudal group</i> | Females | |
| | <i>Bicaudal-C</i> | carrying homozygous | |
| | <i>Bicaudal-F</i> | mutations develop double | |

| | | |
|-------------------|--|---------------------------------|
| <i>Bicaudal-D</i> | abdomen embryonic phenotype. <i>Bicaudal-D</i> mutants cause mis-location of nanos. This leads to repression of <i>bicoid</i> and zygotic hb expression. Subsequently cues to orchestrate the development of the abdomen are initiated. This explains the double abdomen phenotype of the <i>Bic-D</i> mutants. | Wharton and Struhl (1989, 1990) |
|-------------------|--|---------------------------------|

| | | | |
|------------------|----------------------|-----------------------|----------------------|
| c)terminal group | <i>torso</i> | Mutants lack terminal | in: |
| | <i>trunk</i> | anterior | Lehmann (1988) |
| | <i>torsolike</i> | and posterior | |
| | | (telson) | Sprenger |
| | | structures | <i>et al.</i> (1989) |
| | <i>fs(1)polehole</i> | | |
| | <i>(1)Nasrat</i> | | |
| | <i>l(1)polehole</i> | | |

B. Mutants Affecting The Formation of Dorsal-Ventral Pattern

| | | | |
|----------------------|-------------------------------|----------------------------|----------------------------|
| a) dorsalising group | <i>dorsal</i> | all cells at blastoderm | Reviewed |
| | <i>easter</i> | differentiate as if they | in Nüsslein-Volhard (1991) |
| | <i>gastrulation-defective</i> | were located in the dorsal | |
| | <i>Mat(3)2</i> | position of the blastoderm | |

| | | | |
|--|-----------------|--|--|
| | <i>Mat(3)4</i> | giving rise to a dorsalised phenotype | |
| | <i>snake</i> | | |
| | <i>nudel</i> | | |
| | <i>tube</i> | | |
| | <i>pipe</i> | | |
| | <i>spatzle</i> | | |
| | <i>pelle</i> | | |
| | <i>fs(1)K10</i> | dorsalisation of egg and embryo | Wieschaus et al.(1987) Haenlin et al.(1987) |
| b) ventralising group | <i>gurken</i> | ventralisation of egg shell and embryo | Schüpbach (1987) Price et al., (1989) |
| | <i>torpedo</i> | | |
| c) lateralising group | <i>Toll</i> | lateralisation of embryonic pattern May be responsible for setting up the gradient of dorsal-ventral information specifying the dorsal-ventral pattern. | Anderson (1987) Hashimoto (1988) |
| 5. Abnormalities visible in the cuticle of embryos | <i>rubish</i> | variable head defects and holes | Schüpbach and Wieschaus (1989) |
| 6. Apparently normal larvae that do not hatch | <i>reticent</i> | larvae fail to hatch | Schüpbach and Wieschaus (1989) |

eggs, the *bcd* transcript is strikingly localized at the anterior pole, while its protein forms a concentration gradient over the anterior third of the embryo (Driever and Nüsslein-Volhard, 1988a). Evidence from genetic and cytoplasmic transplantation indicates that there are two maternal genes functioning to localize the *bcd* mRNA. Females homozygous for *exuperantia* or *swallow* mutations produce eggs showing the same phenotype as *bcd* mutant (Frohnhofer and Nüsslein-Volhard, 1986; Berleth *et al.*, 1988). In embryos mutant for both these genes, the *bcd* mRNA is diffused and its protein forms a shallow gradient (Berleth *et al.*, 1988; Driever and Nüsslein-Volhard, 1988b).

The expression of several zygotic genes appears to be dictated by the *bcd* protein, either at the transcription level or at the post-transcriptional level. The *bcd* protein is found in the embryonic nuclei, and contains a homeobox domain (Berleth *et al.*, 1988). It is possible that the *bcd* protein acts directly on the zygotic gene by switching on or off their expression, and this control mechanism appears to be concentration-dependent. In the case of the zygotic gap gene *hunchback*, which is required for the development of anterior segments of the embryo, its transcription is activated by *bcd*. The gene *hunchback* is normally expressed in the anterior third of the embryo, at the beginning of the syncytial blastoderm stage, and is transcribed at a uniform level at a particular *bcd* protein concentration. In *bcd* mutant embryos, the transcription of *hunchback* is abolished (Gaul *et al.*, 1987; Tautz, 1988). The *bcd* product excises a negative control on another zygotic gap gene (*Krüppel*), which is needed for the development of abdominal segments (Jäckle *et al.*, 1988). In situ hybridization experiments show that in the absence of *bcd* activity, the *Krüppel* mRNA band moved anteriorly to the region where the thorax is normally formed (Tautz, 1988). This result indicates that the expression of *Krüppel* is normally suppressed by *bcd* in the anterior region of the embryo.

B. The Posterior Group

The posterior group of maternal effect genes comprises at least 12 genes which function to determine the development of abdominal segments and pole cells. Embryos produced by females homozygous for these mutant genes have similar phenotypes, i.e. several abdominal segments are deleted and, except for *nanos* and *pumilio*, all fail to develop pole cells (Nüsslein-Volhard *et al.*, 1987). Like the anterior group gene *bcd*, most mutant embryos in the posterior group can be rescued by transplantation of the pole-plasm from freshly-laid wildtype embryos; in addition, most of the abdominal phenotypes can be rescued by transplantation of the cytoplasm of nurse cells, even from mutant mothers, except in the case of *nanos*. The lack of rescuing ability of the mutant *nanos* indicates that this gene plays a central role in abdomen formation (Lehmann and Nüsslein-Volhard, 1991). Most of the posterior group genes are required to localize the *nanos* product in the posterior pole-plasm, and form the polar granules which contain the determinants for pole cell formation. In mutant embryos of five genes in this group (*oskar*, *staufer*, *tudor*, *valois* and *vasa*), the pole-plasm is eliminated, the localization of *e* plays a central role in abdomen formation (Lehmann and Nüsslein-Volhard, 1991). Most of the posterior group genes are required to localize the *nanos* product in the posterior pole plasm, and form the polar granules which contain the determinants for pole cell formation. In mutant embryos of five genes (*oskar*, *staufer*, *tudor*, *valois* and *vasa*), the pole plasm is eliminated, the localization of *nanos* mRNA is disrupted, and no pole cells are formed (Schupbach and Wieschaus, 1986a,b; Lehmann and Nüsslein-Volhard, 1991).

In controlling posterior pattern formation, the posterior group genes function in a manner different from the anterior group. Unlike the *bcd* gene, which directly activates the transcription of *hunchback*, *nanos* acts indirectly on the zygotic gap-gene *knirps*, which is required for defining the abdomen domain. The gene *knirps* is suppressed by the product of the maternally-expressed *hunchback* gene, which is found uniformly distributed in the freshly laid eggs. The *nanos* product removes the *hunchback*

protein, and thus activates the transcription of *knirps* (Irish *et al.* 1989).

C. Terminal Group

As stated previously, the three gene groups which coordinate the anterior-posterior embryo pattern formation are largely independent from each other in their action; in the case of the double mutant of *bcd* and *oskar*, the telson can still be observed at both ends of the embryo in mirror-image symmetry (Nüsslein-Volhard, 1991). The existence of the telson structure implies that there is a third activity existing in the anterior-posterior system which functions to define the most anterior and posterior structures of the embryo. This activity is provided by the genes of the terminal groups. Mutations of the terminal group genes usually delete the anterior-most and posterior-most unsegmented structures, while the segmented region expands to the both ends of the embryos (Schüpbach and Wieschaus, 1986a). The gene *torso* seems to play a central role in the terminal group. Evidence comes from the existence of dominant gain-of-function alleles of this gene, which produce the phenotype complementary to the lack-of-function alleles (Klingler *et al.*, 1988). The signal produced by *torso* is taken up by the zygotic terminal gene *tailless*, which produces a similar phenotype to the lack-of-function of *torso* (Welgel *et al.*, 1990). Interestingly, the *torso* transcript and protein is uniformly distributed in newly-laid eggs. Molecular studies show that the *torso* gene encodes a putative receptor tyrosine kinase (Sprenger *et al.*, 1987); and its extracellular domain receives the signal generated by another terminal group gene *torso-like*, which is expressed in groups of follicle cells situated at both the anterior and posterior ends of the oocyte during the oogenesis (Stenens *et al.*, 1990).

D. Dorsal-Ventral Group

Twelve maternally-active genes have been identified that are required to coordinate the dorsal-ventral pattern of the embryo. Homozygous mutant females of 11 of these genes produce morphologically normal eggs, but their embryos display a

completely dorsalized phenotype. The only structure that can be found is that normally formed in the dorsal-most region of the embryos, which develops at all positions along the dorsal-ventral axis of the mutant embryo (Nüsslein-Volhard, 1991). Weak alleles of most of these genes have been isolated which produce partially dorsalized embryos with only the ventral-most structures missing (Anderson and Nüsslein-Volhard, 1986).

The continuous spectrum of common phenotypes observed along the dorsal-ventral axis suggests the existence of a concentration gradient of a morphogen, with its highest concentration needed to promote the development of the ventral structures, and intermediate concentrations needed to allow the production of the lateral structures. The cells read the local concentration of this morphogen, and determine their fate during subsequent differentiation. In partial loss-of-function mutants, the total concentration of the morphogen is reduced; thus the morphogenetic signal the cells receive directs them to develop more dorsal like structures. Identification of the gain-of-function alleles of the gene *Toll* (Anderson *et al.* 1985b) has been very helpful in the understanding of the order in which the genes of this group act. The dominant alleles of *Toll* produce ventralized embryos complementary to its recessive alleles and other dorsaling loci, suggesting that the activity of *Toll* is sufficient for the differentiation of the ventral structures. Double mutants for the gain of function allele *Toll^D* and the dorsaling loci *gastrulation defective*, *mudel*, *pipe*, *snake* or *easter* produce lateralized embryos with all the embryonic cells showing lateral structures. The results suggest none of these five genes are essential for the production of the morphogen, and these genes may act upstream of *Toll*. In contrast, the embryos produced by the double-mutant females for *Toll^D*, dorsal or tube display dorsalized structures indicating that the two genes act down stream of *Toll*, because the overly active Toll product is unable to prevent the dorsalized phenotype (Anderson *et al.*, 1985b).

Evidence from molecular studies indicates that the *dorsal* gene is responsible for producing this morphogen (Santamaria and

Nüsslein-Volhard, 1983; Anderson *et al.* 1985b). Unlike the *bcd* product, its protein forms a concentration gradient along the anterior-posterior axis; both *dorsal* mRNA and protein are uniformly distributed in freshly-laid eggs (Roth *et al.*, 1989; Rushlow *et al.*, 1989; Steward, 1989). As soon as the nuclei of the syncytial embryo reach the periphery of the egg, a concentration gradient of *dorsal* protein is formed in the nuclei, with its highest concentration at the ventral side of the embryo (Roth *et al.*, 1989; Rushlow *et al.*, 1989; Steward, 1989). The homology of *dorsal* protein to the oncogene *rel* (Steward, 1989) and the transcription factor NF κ B (Kieran *et al.* 1990, Ghosh *et al.*, 1990), implies that the *dorsal* protein functions as a transcription factor. It has been shown that the expression of zygotic genes *twist* and *zen* are dependent on the local nuclear concentration of *dorsal* protein (Roth *et al.*, 1989; Rushlow *et al.*, 1989). The two zygotic genes are expressed longitudinally along the middle line of the dorsal (*zen*) and ventral (*twist*) sides of the embryo (Rushlow *et al.*, 1987; Thisse *et al.*, 1988). It appears that *dorsal* protein activates the transcription of *twist*, whilst suppressing the expression of *zen* (Doyle *et al.*, 1989).

The gene *Toll* plays a central role in determining the dorsal-ventral polarity. Cytoplasmic injection experiments have shown that among the seven rescuable mutations in the dorsal-group genes (*tube*, *snake*, *easter*, *pelle*, *Toll*, *dorsal* and *cactus*), cytoplasm from mutants in one gene can rescue all other mutant genes as efficiently as wild-type cytoplasm. The rescuing activities of all the loci (except *Toll*) appear to be independent of the site of injection. In other words, the rescued ventral structures are developed from the cells lying on the presumptively ventral side of the egg, irrespective of where the cytoplasm is injected (Anderson and Nüsslein-Volhard, 1986). These results suggest that there is a pre-existing polarity in the embryo, which coordinates the functions of these genes. In contrast, ventral structures are induced at the sites of injection, even at the presumptive dorsal side of the embryo. The purified *Toll* mRNA is used to rescue embryos produced by *Toll*- mothers (Anderson *et al.*, 1985). The ability to induce ventral

structures at ectopic positions by *Toll* transcript leads to the speculation that a local high concentration of *Toll* product is required to establish the polarity, and in wildtype embryos the local activity of *Toll* defines the ventral-most structure. However, molecular analysis has shown that both the mRNA and protein of *Toll* are homogenously distributed along the dorsal-ventral axis of freshly-laid eggs (Gerttula *et al.*, 1988). A possible explanation for this discrepancy is that the *Toll* protein is selectively activated by the gene acting upstream of *Toll*, whose activity has been imprinted by the asymmetry of the egg chamber or mature eggs during oogenesis.

1.5 P-element Mediated Enhancer Detection

1.5.1 P-element Mediated Gene Transfer

A revolutionary development in studying gene function and regulation, as well as identifying unknown genes in *Drosophila melanogaster*, has come from the P-element mediated gene transfer technique (Rubin and Spradling, 1982; Spradling and Rubin, 1982). P-elements are a family of mobile elements originally found in certain *Drosophila* strains called P-strains. Flies which do not have the P-elements in the genome are called M-strains. Crosses between P-strain males and M-strain females frequently induce sterility, and a high rate of recombination and chromosomal aberrations. This phenomenon is called hybrid dysgenesis (Kidwell *et al.*, 1977; Kidwell, 1979). Molecular analysis revealed that the mutations were caused by the insertions and excisions of the P-elements in the genome of *Drosophila*. A number of P-elements have been cloned and sequenced, and these experiments show that although different in size, they all have a conserved sequence at the two terminal repeats (O'Hare and Rubin, 1983; Rubin *et al.*, 1982). A 2.9kb P-element, which encodes the transposase, was found to have the ability to transpose itself and other P-elements (Karens and Rubin, 1984). Rubin and Spradling successfully introduced the *rosy* (*ry*) gene into the *Drosophila* genome to rescue the *ry*-phenotype by the P-element-mediated gene transfer technique (Rubin and Spradling, 1982). A P-transposon containing the

wildtype *ry* gene was injected into *ry*-mutant embryos, together with the 2.9kb wildtype P-element to provide the transposase. Up to 39% of the progeny obtained were *ry*⁺ in the G1 generation, and the *ry*⁺ phenotype was inherited as a stable dominant marker. In this experiment, a wild type P-element had to be co-injected to provide the transposase. The disadvantage of this technique is that while transposing the target transposon, the wildtype P-element itself can be transposed into the fly genome. A major improvement of this technique came from the construction of the P-helper plasmid p π 25.7wc (Karess and Rubin, 1984). Twenty three base-pair nucleotides were removed from the right hand 31bp terminal repeat of the P-element (which is essential for the transposition) in the plasmid p π 25.7 containing a 2.9kb intact P-element. This new plasmid, when injected into *Drosophila* embryos, can produce the transposase but loses the ability to integrate itself into the chromosomes.

1.5.2 Enhancer Trapping

P-element mediated gene transfer is now being used widely to isolate and characterize developmentally regulated genes in *Drosophila*. A major achievement has come from the development of the "enhancer trap" technique (O'Kane and Gehring, 1987). A plasmid called placA92 carrying a P-transposon P[*lac*, *ry*⁺]A (Figure 1.5) was constructed, which contains an in-frame fusion of the *E. coli lacZ* gene to the second exon of the P-element, followed by a genetic marker (*Drosophila rosy* gene). This transposon is randomly integrated into the *Drosophila* genome after injection into *Drosophila* embryos. Due to the weak promoter (P-promoter), the increased expression of the reporter gene should reflect the activity of the enhancers of nearby genes. By staining flies for the presence of β -galactosidase activity, the enhancer elements can be detected in situ by investigating the staining patterns, without knowing the function of the genes. If any interesting staining patterns are found, the transposon can be used as a tag to clone the flanking genomic sequences. This technique has been providing a rapid way of cloning unknown genes with potentially important functions

Figure 1.5

Structure of the plasmid placA92 containing the enhancer detector P[lac, ry⁺]A (from O'Kane and Gehring, 1987)

The plasmid is depicted as linear, with the 5' end of the P-element at the left. Restriction endonuclease sites:

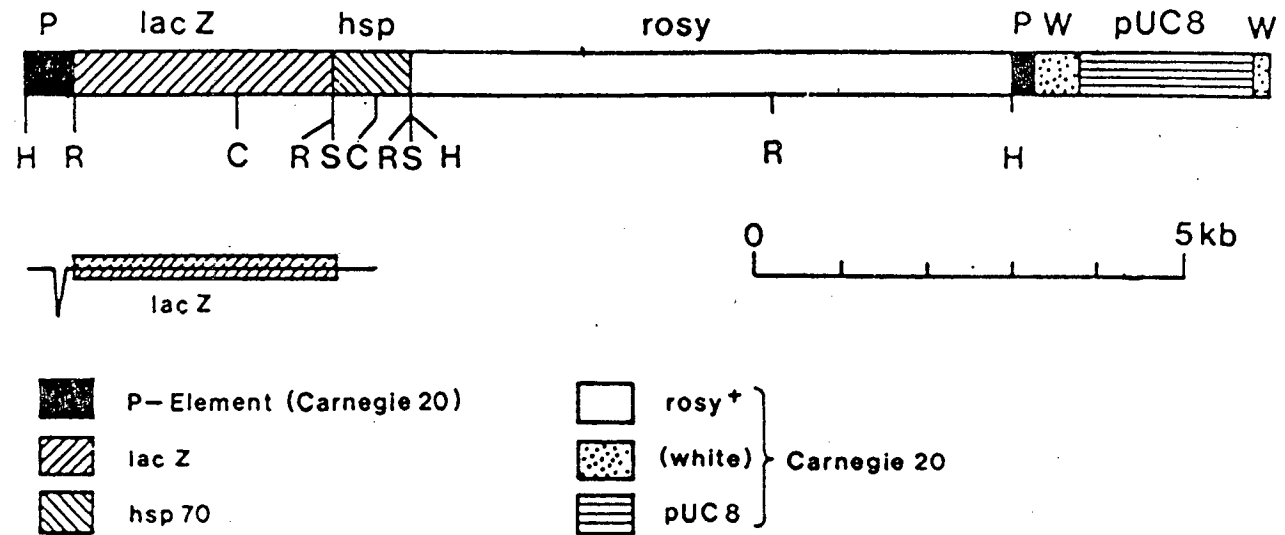
H = HindIII

R = EcoRI

S = SalI

C = SacI

FIGURE 1.5



during development. Cooly and his colleagues further improved this technique by using a *Drosophila* line carrying a special transposon P[ry⁺, Δ 2-3](99B) which functions as the source of transposase but the transposition of the element itself is suppressed (Cooly *et al.*, 1988; Robertson *et al.*, 1988). When the P[ry⁺, Δ 2-3](99B) is introduced into the genome of flies carrying other P-elements by crosses, the P[ry⁺, Δ 2-3](99B) mobilizes other P-elements into new chromosomal positions. This greatly simplified the procedure to generate new "enhancer trap" lines, which used to involve injecting large numbers of embryos.

Several experiments analyzing the *lacZ* staining patterns, either in embryos, or in adults, give strong support to the validity and usefulness of this new technique (Fasano and Kerridge, 1988; Bier *et al.*, 1989; Bellen *et al.*, 1989; Wilson *et al.*, 1989; Grossniklaus *et al.*, 1989; Bownes, 1990)

3768 independent lines carrying the single insertion of PlacW (this transposon contains the P-*lacZ* fusion, the *Drosophila* white gene, and a bacterial origin of replication, as well as the α -lactamase gene) were examined for *lacZ* expression patterns in the embryos (Bier *et al.*, 1989). The results show that 64% of these lines expressed the *lacZ* fusion in the embryos, and most of them show an ubiquitous staining pattern, and the majority of these lines have very specific expression patterns. A similar experiment on a relatively small scale shows similar results (Bellen *et al.*, 1989). 600 independent lines carrying a single transposon P[lArB] (the structure of this transposon is basically the same as P-lacW, except it has the ry gene as the genetic marker) were screened, and 65% of them were claimed to display *lacZ* staining in different cell types in the embryos. From the results of both experiments, the *lacZ* staining lines have been found which mark virtually every known tissue type. In addition, some cell types which are unidentifiable by conventional methods were also stained. The specificity and diversity of staining patterns obviously reflects activities of various endogenous regulatory elements. This is most obvious in the nervous system as results from both experiments show a high proportion of lines with β -galactosidase detected in the nervous

system(49% and 68%, respectively). In most of these lines, only certain cells show the staining in certain cell types or subpopulations of cells. Independent lines were recovered which mark essentially every cell type in the peripheral nervous system (Bier *et al.*, 1989). The high frequency and complex staining patterns in the nervous system may simply indicate the great complexity of the regulatory mechanisms of gene expression during neurogenesis.

The most convincing evidence to support the fundamental assumption of this technique is the good correspondence between the expression patterns of the *lacZ* gene and the nearby endogenous genes. Bier *et al.* (1989) reported several cases where the expression patterns of β -galactosidase and some known genes are well correlated. Bellen *et al.* (1989) mapped the chromosomal location of the P-inserts of the transposants, and showed that their *lacZ* expression patterns are identical to the genes known at the same chromosomal positions.

Fasano and Kerridge (1988) first applied this technique to study oogenesis. 184 lines transformed with P[*lac*, *ry*⁺] were checked for the β -galactosidase expression patterns. A large portion of these lines showed either temporally or spatially specific staining in the ovaries. Grossniklaus *et al.*(1989) carried out more extensive studies with 600 lines of flies carrying a single P[*lArB*] insertion stained for *lacZ* activity. Apart from the diverse staining patterns reflecting the activities of different regulatory elements as discussed earlier for embryonic staining, the results obtained in their experiments have provided several compelling facts which support the credibility of applying this technique to study oogenesis. Based on genetic studies, it has been estimated that 50% to 75% of the recessive female sterile mutations, and 70% zygotic lethal mutations on the X-chromosome are expressed in the germ-line (Perrimon and Gans, 1983; Perrimon *et al.*, 1986). In the *lacZ* staining experiment, it was found that among the strains showing *lacZ* staining in the ovarian cells, 76% of them express β -galactosidase in nurse cells and the germarium. In addition, Perrimon *et al.* (1986) have proposed that very few genes are

exclusively expressed during oogenesis. This agrees with the *lacZ* staining results, as only 2 out of 587 transposants expressed the *lacZ* fusion gene exclusively in the ovary.

The "enhancer" trap technique has several advantages over the conventional genetic approaches for identifying genes important for oogenesis, as it can reveal the expression patterns of genes without knowing their phenotypes. As described in Section 1.4, by genetic screens, large number of female sterile mutation have identified, and some of them have been well characterized. However, it is usually difficult to define in which type of cells a female sterile gene is functioning, since a majority of these fs mutations do not produce easily interpretable phenotypes. The enhancer trap technique has provided an alternative solution for this problem, as different cell types are well marked by the *lacZ* fusion protein, thus making it easier to identify where these genes are expressed. In addition, many fs genes are not only required during oogenesis, but are also involved in other developmental processes, as stated earlier. Mutations on these pleiotropic genes may cause lethality in early development, thus it is impossible to assign the functions during oogenesis, unless hypomorphic alleles exist which only affect oogenesis. However, such alleles have not been frequently identified in the screens for fs mutations (Perrimon *et al.*, 1986, 1989). Finally, the activity of some genes can be complemented by other genes, so it is difficult to identify mutations of these genes by conventional mutagenesis, unless all the other complementary genes are suppressed. Thus the "enhancer trap" technique can be an efficient complementary approach to the conventional genetics to study oogenesis and other developmental processes.

1.6 Aim of This Project

The aim of this project is to cloning and characterizing gene with potential importance during oogenesis.

Since many genes functioning during oogenesis are also essential for viability of flies, it is usually difficult to access whether or not they are active during oogenesis by conventional genetic

techniques. This project took the advantages of the "enhancer trap" technique (discussed in 1.5.2), by staining flies carrying the enhancer detector P[*lac*, *ry*⁺] for the *lacZ* activity in the ovary, to investigate genes which are active during oogenesis.

A 16.5 kb genomic fragment surrounding the P[*lac*, *ry*⁺] insertion was cloned from the strain D19, which showed the *lacZ* staining in ovarian nurse cells and spermatocyte cysts. This genomic fragment encoded multiple transcripts, among them, a 1.1 kb mRNA which was very enriched in the ovary and testis. A cDNA corresponding to this transcript was isolated from a *Drosophila* ovary cDNA library. Both the genomic and cDNA clones were sequenced. The protein sequence predicted from this cDNA showed a high similarity to the sequences of human ubiquitin carboxyl hydrolase isozyme 1 and isozyme 3. The expression pattern of this gene was analyzed using Northern blots and in situ hybridization.

Attempts were also made to mutate this gene by excising the P[*lac*, *ry*⁺] from the genome of D19.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Solutions

Chemicals were obtained from Sigma, BDH and Aldrich.

Enzymes (Restriction Endonucleases, T4 DNA Ligase, Klenow fragment of DNA Polymerase, DNase I, RNase A, Proteinase K) were obtained from BRL, Pharmacia and NBL.

Radioisotopes were obtained from Amersham.

Hybond-N nylon membranes were obtained from Amersham.

Standard solutions were made using sterile distilled water in sterile, baked glassware and were generally made sterile by autoclaving (15psi/15 min). Solutions not detailed in the text are described below.

Boehringer Mannheim Incubation Buffers for Restriction Enzymes :

Final concentrations in mM/L (i.e. 1:10 diluted buffer).

- A: 33mM Tris HCl, 10mM Mg-Acetate, 66mM K-Acetate, 0.5mM Dithiothreitol (DTT), pH 7.5
B: 10mM Tris HCl, 5mM MgCl₂, 100mM NaCl, 1mM 2-Mercaptoethanol, pH 7.5
L: 10mM Tris HCl, 5mM MgCl₂, 1mM Dithioerythritol (DTE), pH 7.5
M: 10mM Tris HCl, 10mM MgCl₂, 50mM NaCl, 1mM DTE, pH 7.5
H: 50mM Tris HCl, 10mM MgCl₂, 100mM NaCl, pH 7.5

| | |
|-------------------------------------|--|
| 0.5M EDTA | 0.5M Diaminoethanetetra - acetic acid, pH8.0 |
| 10X TBE gel buffer: | 89M Tris-Borate, 0.89M Boric Acid, 10mM EDTA |
| 10X DNA and RNA gel loading buffer: | 100 mM EDTA pH 8.0, 0.1% (w/v) Bromophenol Blue, 20% (w/v) Ficoll (Type 400) |
| 10X MOPS gel buffer: | 2M Na-MOPS pH 7.0, 50mM Na-Acetate, 10mM EDTA |
| RNA extraction buffer: | 100mM Tris HCl pH 7.5, 10mM EDTA, 150mM Lithium Chloride, 1% (w/v) SDS |

OLB

(oligo-labelling buffer):

| | |
|-------------------------|---|
| Solution 0: | 0.125M MgCl_2 , 1.25M Tris HCl pH 8.0 |
| Solution A: | 0.95ml solution 0, 18ml 2-Mercaptoethanol, 25ul 20mM dATP, 25ul 20mM dTTP, 25ul 20mM dGTP |
| Solution B : | 2M Hepes, pH 6.6 |
| Solution C: | Hexadeoxyribonucleotides (Pharmacia) suspended in TE buffer at 90 OD units/ml |
| | OLB is made by mixing Solution A, B and C in the ratio 2:5:3 (and is stored at -20C) |
| 20X SSC: | 3M NaCl, 0.3M Na- Citrate |
| Salmon Sperm DNA: | 10mg/ml stock which has been sonicated, phenol extracte and ethanol precipitated. Purchased from SIGMA. |
| Hybridization Solution: | 50% (v/v) formamide, 7% (w/v) SDS, 0.4 M $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ pH 7.0 |
| Ringer's Solution: | 3.2g NaCl, 3.0g KCl, 1.8g MgSO_4 , 0.69g $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 1.79g Tricine, 3.6g glucose, 17.1g sucrose. Made up to 1 litre with dH_2O , pH adjusted to 6.95, filter sterilised and stored at 4C |
| 10X PBS: | 0.1M Sodium Phosphate pH 7.5, 1.3M NaCl |
| TM buffer: | 10mM Tris HCl pH 8.0, 10mM MgCl_2 |
| TE buffer: | 10mM Tris HCl pH 7.4-8.0, 1mM EDTA |
| DEPC treated water: | DEPC was added to dH_2O to 0.1% (v/v), and left at 37°C overnight, then autoclaved for 15 minutes. |

2.2 *Drosophila* Stocks and Strains

Drosophila strains are listed in Lindsley and Grell (1968) and Lindsley and Zimm (1987 and 1990). Table 2.1 presents a list of stocks used.

Table 2.1 List of *Drosophila* Strains and Stocks Used

| <u>Stocks</u> | <u>Relevant Features</u> | <u>References</u> |
|---------------|--------------------------|----------------------------------|
| Oregon R | A wildtype strain | Lindsley and Grell (1968) |
| E#2 | A P-transposase source. | Robertson <i>et al.</i> , (1988) |

| <u>Balancer Chromosomes</u> | <u>References</u> |
|-----------------------------|-------------------|
|-----------------------------|-------------------|

| | |
|-----|---------------------------|
| CyO | Lindsley and Grell (1968) |
| TM6 | |

P[lac, ry[±]]A transposant strains

All the P[lac, ry⁺]A transposant strains are gift from Prof. John Merriem, University of California. These strains carrying the P[lac, ry⁺]A insertion on chromosome 2 or chromosome 3.

Table 2.2: Stage of *Drosophila melanogaster* development at 25°C
(Reference: Bownes and Dale)

| <u>Hours</u> | <u>Days</u> | <u>Developmental stages</u> |
|--------------|-------------|--|
| 0 | 0 | Fertilization and fusion of pronuclei. |
| 1.5 | | Preblastoderm stage. Migration of cleavage nuclei and pole cell formation. |
| 3 | | Blastoderm stage. Migration of nuclei form cells in the previously syncytial blastoderm. |
| 3.5 | | Gastrulation begins. |
| 6-8 | | Segmentation visible. |
| 18 | | Larval differentiation nearly complete. |
| 24 | 1 | Hatching from egg. Onset of first larval instar. |
| 48 | 2 | First moult. Second larval instar begins. |
| 72 | 3 | Second moult. Second larval instar begins. |
| 120 | 5 | Prepupal moult. Pupation. Eversion of imaginal discs |
| 122 | | Puparium formation with white puparium. |
| 132 | | Puparium darkens |
| 216-240 | 9-10 | Emergence of adult from pupal case. |

Drosophila stocks were maintained in vials or bottles at 18°C or 25°C on cornmeal food which consisted of cornflour 250 g, sucrose 500g, yeast pellets 175 g and agar 100g made up to 10 litres with dH₂O. The food was then boiled, cooled to approximately 40°C and poured into bottles or vials and allowed to set. A fungicide, Nipagen was added to a final concentration of 4.5 ug/L and on occasion antibiotics such as Gentamycin (to 40 ug/ml) were added. Strips of filter paper that had been soaked previously in 3% (v/v) Benzyl Benzoate in ethanol and allowed to air dry, were placed on top of the food if mite infections occurred.

2.3 *Drosophila melanogaster* Life Cycle - Collection of Developmental Stages

2.3.1 Stages of *Drosophila melanogaster* Development at 25°C

Table 2.2 summarises the stages of *Drosophila melanogaster* development at 25°C for the strain used, Oregon R.

2.3.2 Collection of Staged Eggs

500ml plastic beakers were used as population cages. Flies were raised in milk bottles for 3-4 days after eclosion before being transferred into the plastic beakers. About 300-500 flies were transferred into individual beaker. The mouth of the beaker was then covered with Petri dish containing the ADH fly food with the yeast paste smeared in the middle. Four cages were set each time. Eggs were washed off from the Petri dish at two hour intervals and collected in a sieve and rinsed thoroughly using distilled water. The eggs were then subject to subsequent treatment or frozen in liquid nitrogen and stored at -70°C until use.

2.3.3 Collection of Early and Late 3rd Instar Larvae and Pupae

Egg laying flies were placed in fresh food bottles for 4 hours then removed. The bottle was then incubated for 72 hours, 96

hours and 120-240 hours. Early 3rd instar larvae were collected after 72 hours by floating them out of the food using a saturated sucrose solution, rinsed with dH₂O and frozen in liquid nitrogen. Late 3rd instar larvae were picked with a paintbrush as they crawled up the sides of the bottles after an incubation period of 96 hours, placed in microcentrifuge tubes and frozen with liquid nitrogen, and stored at -70°C. Larvae were sexed according to the size of their gonads, males have much bigger gonads than females.

Pupae were collected by picking them off the inside of the food bottles after 120-140 hours incubation, placed in microcentrifuge tubes and frozen in liquid nitrogen, and stored at -70°C.

2.3.4 Collection of Sexed Adults

Adult flies were collected ranging in age from newly eclosed to 10 day old flies. The flies were anaesthetized with diethyl ether, sexed, placed in microcentrifuge tubes, frozen in liquid nitrogen and stored at -70°C.

2.4 Microbial Strains and Media

2.4.1 Microbial Strains

All bacterial strains, plasmid vectors, bacteriophage, plasmids and libraries used are listed in Table 2.3.

Bacterial stocks were maintained on appropriate plates at 4°C and also as stabs at room temperature in the dark.

Lambda bacteriophage were stored at 4°C in phage buffer with a few drops of chloroform added.

Plasmid and Lambda DNA samples were stored suspended in TE buffer and kept at -20°C.

2.4.2 Media

All media used (listed in Table 2.4) were sterilised by autoclaving. Antibiotics, vitamins and sugars were made up in dH₂O and filter sterilised. Where appropriate, antibiotics were added to plates and media, e.g. ampicillin to a final concentration of 50 µg/ml.

2.5 Methods

2.5.1 General Methods

2.5.1.1 Phenol Extraction

Water saturated phenol was equilibrated with TE buffer and 8-hydroxyquinoline was added to 0.1% (w/v). Solutions of nucleic acid were deproteinized by shaking, or vortexing, with an equal volume of phenol/chloroform (1:1). The phases were then separated by centrifugation and the aqueous layer were transferred to a fresh container.

2.5.1.2 Precipitation of Nucleic Acids

The solution was adjusted to 0.3 M Sodium Acetate pH 5.5 then 2.5 volumes of ethanol or 0.6 volumes of isopropanol were added.

The samples were incubated at -20°C for a minimum of 60 minutes. Small quantities of nucleic acids (if less than $5\mu\text{g}$) were precipitated at -20°C overnight. The precipitate was recovered by centrifugation in a Sorval centrifuge (10 Krpm, 20 minutes, 4°C) or a microfuge (12 Krpm, 10 minutes). The pellet was washed in 70% (v/v) ethanol, dried under vacuum and re-dissolved in TE buffer or sterile, distilled water.

2.5.1.3 Restriction Endonuclease Digestion of DNA

DNA was digested with a 2-4 fold excess of the enzyme for at least 60 minutes in conditions recommended by the manufacturer. After digestion, samples were either directly loaded for electrophoresis, or deproteinized by phenol extraction.

2.5.1.4 Extraction of DNA from Agarose gels (GENECLEAN KIT, BIO 101 Inc)

Agarose gels were run until the fragments were well separated and the gel slice containing the desired fragment was cut out. The gel slice was transferred into a microcentrifuge tube, to it, 4.5 volumes of NaI and 0.5 volumes of TBE modifier were added. The tube was incubated at 45°C for 5 minutes until the gel slice was melted. The $5\mu\text{l}$ glass milk was added and mixed thoroughly, and

Table 2.3 Microbial Strains, Vectors and Libraries Used

| <u>Host</u> | <u>relevant genotype</u> | <u>Comments</u> | <u>Reference</u> |
|------------------------|--|--|---------------------------------|
| G600 | supE | used for growing λ NM1149 and λ Dash | Jendrisak et al.(1978) |
| NM514 | hfl hsdR | selection for cl^{-} - recombinat | Murray (1983) |
| NM522 | supE, thi, hsd5, (lac-proAB), [FproAi3, lacI ^q ZM15] | Used as a host for pBluescript and pGEX-2T | Cough & Murray (1983) |
| Y1090 | hsdR, supF Δ lac hsdR Δ lon pMC9 | Used as a host for λ gt11 | Young & Davis (1983) |
| <u>Phage</u> | | | |
| λ NM1149 | | immunity insertion vector, for EcoRI and HindIII fragment vector, for EcoRI and HindIII fragment | Murray (1983) |
| VCM13 | Kanamycin ^r | Used for ssDNA rescuing | |
| <u>Plasmids</u> | | | |
| pGem1-lacZ α | Amp ^r | pGem1 containing 3 kb BamHI lacZ fragment | Bownes & Shirras (1987) |
| pGem Adh | Amp ^r | pGem1 containing 1.7 kb XbaI fragment of Drosophila Adh gene | |
| pGem α -tubulin | Amp ^r | pGem2 containing 1.7 kb α -tubulin-1 DNA fragment | Kalfayan & Wensink (1982) |

| | | | |
|----------------|------------------|--|------------|
| pBluescriptSK+ | Amp ^r | A phagemid vector for subcloning, producing ssDNA in vitro RNA transcription | Stratagene |
| pGEX-2T | Amp ^r | A procaryotic gene fusion vector for expressing fusion proteins | Pharmacia |

| <u>Libraries</u> | <u>Comments</u> | <u>Source</u> |
|---|--|---|
| λDash Drosophila Genomic library | Sau3A partial digested Drosophila DNA and cloned into BamHI site of λDash I vector | This library was constructed by RSauders, University of Dundee |
| λgt11 Ovarian cDNA Library | Whole ovaries were hand dissected from Canton S strain of Drosophila melanogaster and was constructed and poly (A) ⁺ RNA prepared. cDNA was made by oligo (dT) priming and packaged into λgt11 vector | Huynh, T. et al, (1985) (The library donated by P. Sullivan and L. Kalfaya, University of North Carolina |

Table 2.4: Media

| | |
|----------------------|--|
| L-broth Extract, | 10g Difco Bacto Tryptone, 5g Difco Bacto Yeast 5g NaCl per litre, pH 7.2 |
| L agar Extract, | 16g Difco Bacto Tryptone, 10g Difco Bacto Yeast 5g NaCl per litre |
| BBL agar 5g NaCl, | 10g Baltimore Biological Laboratories trypticase, 10g Difco agar per litre |
| BBL top agar | As BBL agar except, only 6.5g Difco agar per litre |
| Top agarose | 0.7% (w/v) Pharmacia agarose in L-broth |
| Phage buffer | 3g KH_2PO_4 , 7g Na_2HPO_4 , 5g NaCl, 10ml 0.1M MgSO_4 , 10ml 0.1M CaCl_2 , 1ml 1% (w/v) gelatin per litre |

the tube was incubated on ice for 5 minutes. After DNA bound to the silica matrix, the silica matrix was pelleted by a brief centrifugation. Then the silica matrix was washed three times each for 5 minutes using the ice cold "New Wash" supplied in the kit. DNA was eluted from the silica matrix by adding appropriate volume of TE buffer and incubating at 55°C for 5 minute. The silica matrix was pelleted by centrifuging for 30 seconds, and the supernatant containing DNA was transferred into a fresh tube.

2.5.1.5 Ligation of DNA

The vector and insert DNA were cut to completion with appropriate restriction endonucleases. After restriction and removing the restriction endonucleases, DNA was ethanol precipitated. Typically, between 100-200 ng insert DNA were ligated in a reaction with vector to insert DNA concentration at a 1:1 molar ratio. Ligations were carried out in 10 μ l reaction volume containing 20 mM Tris-Cl pH 7.6, 5 mM MgCl₂, 5 mM DTT and 1 mM ATP. and incubated overnight at 15°C (for cohesive ends), or room temperature (for blunt ends). 1 unit of T4 DNA ligase was used for blunt end ligations, and 0.1 units for cohesive end ligations. Ligation product were then transformed into *E. coli* or packaged in lambda packaging extracts.

2.5.1.6 Preparation of Dialysis Tubing

Dialysis tubing (32 mm flat width) was purchased from SIGMA. The tubing was cut into pieces of convenient length (10-20 cm), and boiled for 10 minutes in 2 litres of 2% (w/v) sodium bicarbonate/1mM EDTA. The tubing was then thoroughly rinsed in dH₂O followed by boiling in dH₂O for 10 minutes, allowed to cool and stored submerged in 70% (w/v) ethanol at 4°C. Prior to use the tubing was rinsed in dH₂O.

2.5.1.7 Transformation into *E. coli* (Chung *et al.*, 1989)

100 ml of L-broth was inoculated with 1 ml of an overnight culture of relevant bacterial strain, and grown to OD₆₀₀ = 0.3, at 37°C with shaking. The cells were then pelleted by centrifugation for 5 minutes at 4°C (5000 rpm, in the GSA34 rotor), resuspended

in 10 ml ice-cold TSS(L-broth with 10% (w/v) PEG600, 5% (v/v) DMSO), 40 mM MgCl₂, pH6.5). One 100 μ l aliquot was removed for each transformation, mixed with the DNA and incubated on ice for 40 minutes, before spreading, after 1 hour incubation for antibiotic resistance expression. X-gal and IPTG were added as required.

2.5.2 Detection of β -galactosidase in Adult Flies

Adult flies and larvae were dissected in Ringers' solution. Ovaries, testes and fly carcasses were separately transferred into the wells of a microtite plate containing 100 μ l staining solution(see below). Staining was conducted at room temperature for over night or at 37°C for 2 hours. After staining, tissues were transferred onto a microscope slide and mounted in Ringers' solution. The specimen were observed with a microscope and photographed.

Staining solution:

| | |
|------|--|
| 1 ml | 10 mM NaH ₂ PO ₄ /Na ₂ HPO ₄ .2H ₂ O pH 7.2 |
| | 150 mM NaCl |
| | 1 mM MgCl ₂ .6H ₂ O |
| | 7 mM K ₄ [Fe ^I (CN) ₆] |
| | 7 mM K ₄ [Fe ^{II} (CN) ₆] |
| | 0.3% (v/v) Triton X-100 |

Just before use, 25 μ l 8% (w/v) X-Gal (in dimethylformamide) was added

2.5.3 Preparation of Genomic DNA from Adult flies

The method of Jowett (1986) was used with some modification. 10 flies were etherized, frozen in liquid nitrogen and homogenized in 350 μ l homogenizing buffer (0.15 M NaCl, 50 mM Tris-HCl pH 8.0, 15 mM EDTA). After homogenization another 350 μ l above buffer with 0.04% (w/v) SDS was added. Proteinase K was supplemented to a final concentration of 50 μ g/ml. The sample was incubated for 1 hour at 65°C. After incubation, the homogenate was extracted twice with phenol/chloroform (1:1, v/v), and once with chloroform. DNA was precipitated with isopropanol. The DNA pellet was washed

with 70% ice cold ethanol, dry briefly under vacuum and dissolved in 45 μ l TE or dH₂O.

2.5.4 Preparation of RNA from Different Developmental Stages and Tissues of *D. melanogaster*

Tissues from adult flies were hand-dissected in Ringers' solution. Methods for collection of embryos, larvae and pupae were described in 2.3.2 and 2.3.3, respectively. Tissues were frozen in liquid nitrogen and homogenized in 750 μ l RNA extraction buffer. The homogenate was extracted 3 times with 1:1 ratio of phenol:chloroform and then isopropanol precipitated for at least 4 hours at -20°C. The pellet was washed in 70% (v/v) ice-cold ethanol before being resuspended in dH₂O and 3 volumes of 3 M LiCl added. The sample was incubated at 4°C overnight, centrifuged and the pellet washed on 70% (v/v) ice-cold ethanol, dried and resuspended in DEPC treated water.

2.5.5 Manipulation of Lambda

2.5.5.1 In Vitro Packaging of Phage Lambda DNA (Amersham)

Packaging Extract A and B were removed from the freezer, and allowed to thaw on ice. Immediately after extracts were thawed, DNA (5-10 μ l) to be packaged was added to Extract A followed by transferring 15 μ l Extract B in to the same tube, and mixed gently with the tip of the pipette used for the transfer. The tube was centrifuged for 10 seconds, to collect contents at the bottom. This mixture was incubated at 20°C for 120 minutes. 0.5 ml phage buffer was added and followed immediately by 10 μ l chloroform, mixed gently.

2.5.5.2 Preparation of lambda DNA

a) Mini-scale

Single plaques were picked from BBL plate and transferred into a 1.5 ml microcentrifuge tube containing 1 ml phage buffer with few drops of chloroform and left at 4°C for at least 6 hours. 1.5 ml fresh over-night culture of appropriate host cells was spun

down and resuspended in 1 ml 10 mM MgCl_2 . 200 μl above phage suspension was added to the cells and incubated for 15 min at 37°C . This cell-phage mixture was added to 10 ml L-broth in a flask with the supplement of MgCl_2 to a final concentration of 10 mM. The flask was incubated at 37°C over night with vigorous shaking. After the cells were lysed, cell debris was removed by centrifugation (7 Krpm). To the supernatant, DNase I and RNase A were added to a final concentration of 1 $\mu\text{l}/\text{ml}$ for both enzymes. After incubation at room temperature for 30 min, phage particles were harvested by ultra-centrifugation for 1 hour (30 Krpm, Sorvall OTD-50B). The phage pellet was resuspended in 400 μl TM buffer. The phage suspension was incubated at 65°C for 1 hour with presence of 1% (v/v) SDS, 0.02 M EDTA and 1 $\mu\text{g}/\text{ml}$ Proteinase K. The phage DNA was extracted twice with phenol/chloroform and once with chloroform, and ethanol precipitated. The phage DNA was dissolved in TE and stored at -20°C .

b) Maxi-scale

Host cells were prepared as described for the mini-scale. 1 ml phage suspension was combined with 2ml host cells and incubated at 37°C for 15 min. The cell-phage mixture was added to 250 ml L-broth in a flask with MgCl_2 added to the final concentration of 10 mM. The culture was shaken at 37°C over night. The lysate was centrifuged at 7 Krpm for 20 minutes to remove cell debris, and the supernatant was transferred into a clean flask. The supernatant was treated with DNase I and RNase A (1 $\mu\text{g}/\text{ml}$ for both enzymes) at room temperature for 30 minutes. Phage particles were precipitated by adding PEG 6000 to a concentration of 10% (w/v) and NaCl to a concentration of 1 M. This was incubated on ice for at least 1 hour. The phage was spun down at 10 kg for 10 minutes, and the pellet was resuspended in 10 ml phage buffer. This phage suspension was extracted once with chloroform. CsCl was added to a concentration of 0.75g/ml and the phage suspension was sealed in a 12 ml Ti50 tube, centrifuged at 30 Krpm (Sorvall Ti50) for 24 hours. After ultra-centrifugation, the phage band was collected through the wall of the tube with a syringe and dialysed against 50

mM Tris-Cl, pH 8.0, 1 mM EDTA at 4°C over night. The dialysed phage suspension was incubated at 65°C for 1 hour with supplement of SDS (1% w/v) and Proteinase K (50 µg/ml). Phage DNA then was deproteinized by phenol/chloroform extraction. DNA was ethanol precipitated and redissolved in 500 µl TE.

2.5.5.3 Plating Lambda Libraries

Genomic libraries were plated on BBL plates. A ovarian cDNA library was plated on BBL-amp plates. Approximately 50,000 recombinant phage were mixed with 1 ml fresh over night culture of appropriate host cells, and incubated for 15 minutes at 37°C. To this 25 ml top BBL-agar (cooled to 46°C) was added and quickly poured onto a pre-dried BBL-bottom agar plate. Once the top agar is set, the plate was inverted and incubated overnight at 37°C.

2.5.5.4 Taking Plaque Lift

After overnight incubation, the plate was allowed to stand for 1 hour at 4°C. Then a piece of Hybond nylon membrane cut into the same size as the plate, was laid on the surface of the top agar for 1 minute, and the orientation was marked with a syringe needle. The membrane was gently lifted and placed, with the DNA side up, on a blotting paper soaked with denaturing solution for 7 minutes. The filter then was neutralized by being placed on blotting papers soaked with neutralizing solution twice, each for 3 minutes. The filter was washed briefly in 2XSSC, blot dried before being fixed under UV light for 10 minutes. Membrane filter was the subjected for hybridization.

2.5.6 Preparation of Plasmid DNA

2.5.6.1 mini-scale (Maniatis, 1978)

Single colony of bacteria containing required plasmid was used to inoculate 5 ml L-broth containing 5mg/ml ampicillin, and the culture was shaken at 37°C overnight. 1.5 ml overnight culture was centrifuged for 1 minute in a microcentrifuge (12 Krpm). The supernatant was discarded, and cell were resuspended in 350 µl GTE (8% sucrose (w/v), 0.5% (v/v) Triton X-100, 50 mM EDTA pH 8.0, 10 mM Tris-Cl pH 8.0) and 25 µl lysozyme (10 µg/ml in 10 mM

Tris-Cl pH 8.0). Cells were lysed by boiling for 40 seconds. Then, cell debris was spun down at 12 Krpm for 10 minutes, and removed using a sterilized toothpick. DNA was precipitated with 40 μ l 2.5 M Na-acetate pH 5.2 and 400 μ l isopropanol. DNA pellet was washed with 70% (v/v) ice-cold ethanol and dissolved in 50 μ l TE.

2.5.6.2 Maxi-scale

1 ml fresh overnight culture was used to inoculate 500 ml L-broth with appropriate antibiotics. Cells were allowed to grow at 37°C overnight with vigorous shaking. Cells were harvested by centrifuging at 5 Krpm for 5 minutes, and resuspended in 10 ml solution I (50 mM Tris-Cl pH 8.0, 10 mM EDTA, 400 μ g/ml RNase A). To this suspension, 10 ml solution II (200 mM NaOH, 1% (w/v) SDS) was added and gently mixed. Cells were allowed to lyse on ice for 5 minutes, and then neutralized with 10 ml 2.55 M Na-Acetate pH 4.5. Cell debris was removed by centrifugation at 10 kg for 30 minutes at 4°C. Supernatant was transferred to a clean centrifugation bottle, and DNA was precipitated with 0.8 volume of isopropanol and spun at 10 kg for 30 minutes at 4°C. After being washed with 70% (v/v) ice-cold ethanol, DNA was dissolved in 10 ml TE, and 10g CsCl and 1 ml ethidium bromide (10 mg/ml) were added. DNA was banded in a Sorvall OTD50B ultracentrifuge at 38 krpm for 48 hours. The supercoiled plasmid band was collected with a syringe under UV light. Ethidium bromide was extracted by water saturated butan-2-ol. DNA was dialysed four times against 1 litre TE.

2.5.7 Electrophoresis

2.5.7.1 Agarose electrophoresis of DNA

Agarose was melted in 1 X TBE buffer, and ethidium bromide was added to a final concentration of 0.5mg/ml. The gel was then casted in a gel former of appropriate sizes.

Concentrations of agarose, and voltage for running gels differ according to purpose. After electrophoresis, DNA was visualized under UV light and a photograph was taken.

2.5.7.2 Agarose gel for RNA

Denature agarose gels were used for electrophoresis of RNA samples. Gels were made up with 1.3-1.5% (w/v) agarose in 1 X MOPS buffer and 17.3% (w/v) formaldehyde. 12.5 μ l formamide, 2.5 μ l 10 X MOPS buffer, and 4 μ l 37% (w/v) formaldehyde were added to 6 μ l RNA samples containing about 10 μ g total RNA. Samples were heated for 5 minutes at 65°C and chilled on ice for a few minutes before 2.5 μ l loading buffer was added. Samples then were loaded onto the gel and run in 1 X MOPS buffer at 100v for 3 to 4 hours.

2.5.7.3 SDS-Polyacrylamide Gel Electrophoresis of Proteins

After IPTG induction, 500 μ l cells were pelleted by centrifugation, and resuspended in 80 μ l loading buffer. Samples then were boiled in a water bath for 5 minutes and centrifuged for 3 minutes before being loaded.

10% SDS-PAGE gel was used. After electrophoresis, gels were fixed in 20% (w/v) TCA for 30 minutes, then stained in 0.1% Coomassie Blue in a 5:5:1 mixture of acetic acid, methanol and dH₂O for 1-2 hours. Gels were destained in a 5:5:1 mixture of acetic acid, methanol and H₂O overnight.

Protein gel solutions:

2X Loading Buffer:

50 mM Tris-Cl pH 6.8; 1% (w/v) SDS, 20% (v/v)
glycerol, 0.1% (v/v) β -mercaptoethanol, 0.01% (w/v)

Bromophenol blue

Separating Gel Solution:

| | |
|------------------------------------|-------------|
| 3 M Tris-Cl pH 8.8 | 5 ml |
| 10% (w/v) SDS | 0.4 ml |
| 30%:0.8% Acrylamide:Bis-Acrylamide | 13.3 ml |
| dH ₂ O | 21.3 ml |
| 15% (w/v) APS | 133 μ l |
| TEMED | 25 μ l |

Stacking Gel Solution:

| | |
|------------------------------------|---------|
| 0.5 M Tris-Cl pH 6.8 | 1.25 ml |
| 30%:0.8% acrylamide:bis-acrylamide | 1.25 ml |

| | |
|-------------------|------------|
| 10% (w/v) SDS | 0.1 ml |
| dH ₂ O | 7.3 ml |
| 15% (w/v) APS | 75 μ l |
| TEMED | 7 μ l |

10X Gel running buffer:

| | | |
|-----------|--------|---|
| Glycine | 144 g | Made up to 1 litre with dH ₂ O |
| Tris-base | 30.2 g | |
| SDS | 10 g | |

2.5.7.4 Polyacrylamide Gels for DNA sequencing

Sequencing gels were 0.6% acrylamide (19:1 acrylamide:bis-acrylamide) in TBE gel buffer containing 7.7M urea. 45 ml above gel mix was used for one gel. 120 μ l APS and TEMED were added and mixed well before pouring the gel.

Sequencing gels were run at a constant power of 40 watts for times ranging between 2.5-14 hours. After electrophoresis, gels were fixed in 10% (v/v) acetic acid, 10% (v/v) methanol for 15-20 minutes, fixed on blotting paper, dried down on a gel drier, and autoradiographed for 16-36 hours at room temperature.

2.5.8 Transfer Nucleic Acid to Membranes

2.5.8.1 Southern Blot (Southern, 1975)

After electrophoresis and photography, gels (containing DNA samples) were merged in denaturing solution for 30 minute with gentle shaking. Then they were soaked in neutralizing solution twice each for 15 minutes. A wick made of wet blotting papers was placed on a platform with both ends of the blotting paper immersed in 20 X SSC in a reservoir underneath the platform. The gel was laid on the top of the wick, and a sheet of Hybond-N membrane (Amersham) cut to the same size as the gel was placed on the top of the gel. The edges were sealed with plastic films. Two sheets of blotting paper soaked in 20X SSC were laid on the top of the membrane, and another two sheets of dry blotting paper were placed on the top of the wet blotting papers followed by a stack of paper towels. A sheet of glass plate was placed on the top of the paper towels, and a weight was laid on the top. A minimum of 2.5

hour was used for transfer. When transfer was complete, the filter was marked with a pencil for the position of the sample wells. Then the membrane was carefully lifted, rinsed briefly in 2X SSC. DNA was fixed to the filter in a UV box (Chromato Vue C-70G, UV view system) with the DNA side facing the UV light source for 10 minutes.

2.5.8.2 Northern Blot (Thomas, 1980)

After electrophoresis, the gel was transferred onto a platform with a blotting paper wick set up in the same way as Southern blot.

As soon as transfer is completed, the filter was UV fixed without washing.

2.5.9 Labelling Nucleic Acid

2.5.9.1 Radioactive Labelling DNA by Random Primering (Feinberg and Vogelstein, 1983, 1984)

DNA fragment to be labelled was denatured by boiling for 10 minutes followed immediately by rapid cooling on ice. Between 50-200ng of DNA was labelled in a reaction in a final volume of 50 μ l containing 10 μ l OLB, 1 μ l 20mg/ml BSA, 20-50 μ Ci 32 P-dCTP, 1 μ l (5 units/ μ l) Klenow fragment of DNA polymerase I. The reaction was incubated for 1-16 hours at 37°C.

2.5.9.2 Digoxigenin Labelling DNA by Random Primering

DNA was denatured as radioactive labelling. Between 1-2 μ g DNA was labelled in a 20 μ l reaction containing 2 μ l primer mix, 2 μ l dNTP, and 5 units of Klenow fragment of DNA polymerase I. The reaction was incubated for 20 hours at 37°C.

2.5.9.3 Digoxigenin Labelling RNA by *In Vitro* Transcription

DNA fragments were cloned into The vector pGem or pBluescript SK⁺ which contain T7 promoter of Bacterial phage. Before transcription, 2 μ g DNA was digested with restriction endonuclease which cuts at one end of the insert distal from the T7 promoter. The digested DNA was phenol extracted, ethanol

precipitated and dissolved in 15 μ l DEC treated dH₂O. To this, 2 μ l NTP (10 mM ATP, CTP, GTP; 6.5 mM UTP; 3.5 mM digoxigenin-11-UTP), 1 μ l RNase inhibitor (20 units/ μ l) and 2 μ l (20 units/ μ l) T7 RNA polymerase were added. The reaction was incubated for 2 hours at 37°C. The labelled RNA was ethanol precipitated with 2 μ l yeast tRNA (50 mg/ml) and 2.5 μ l 4M LiCl.

2.5.10 Hybridization Labelled Probes to Membrane Filters

Filters obtained from Southern bolt, Northern Blot and plaque lift were prehybridized in polythene bags (Krups Vacupac) containing hybridization solution (150 mm²/10ml) at 42°C for at least 1 hour. DNA probes were added to the bags after denaturation by heating at 95°C for 10 minutes. Hybridization was allowed to proceed at 42°C overnight. After hybridization, Southern blots of genomic DNA, and Northern blots were washed twice in 2X SSC, 0.1% (w/v) SDS (w/v) at 65°C for 20 minutes each. Plaque lifts and Southern blots of lambda DNA and plasmid DNA were washed in 0.1XSSC and 0.1% SDS (w/v) for the same time as used for Northern blots. After washing, filters were subjected to signal detection.

2.5.11 Signal Detection for Membrane filters and Sequencing Gels

2.5.11.1 Autoradiography

Autoradiography of dried sequencing gels, membrane filters was performed using Cronex 4 (Dupont) X-ray films and cassettes. Films were developed in a Agfa 1 automatic film processor.

2.5.11.2 Immuno-detection Digoxigenin Labelled Probes

After washing with SSC (2.5.9), filters were rinsed briefly in buffer 1 (100 mM Tris-Cl, 150 mM NaCl; pH 7.5). Then filters were block for 30 minutes in buffer 2 (0.5% (w/v) blocking reagent (supplied in Boehringer's Kit) in buffer 1). After a brief rinse in buffer 1 again, filters were incubated with anti-digoxigenin-antibody-conjugate (1:5000 dilution in buffer 1) for 30 minutes. Unbound antibody-conjugate was removed by washing 2

X 15 minutes in buffer 1. Filters were equilibrated for 2 minutes in buffer 3 (100 mM Tris-Cl, 100 mM NaCl, 50 mM MgCl₂; pH 9.5) and transferred into a polythene bag containing the colour solution (45 μ l NBT solution (75 mg/ml in 70% (v/v) dimethylformamide) and 35 μ l X-phosphate (50 mg/ml in dimethylformamide)). Colour development was carried out in the dark, and stopped in TE buffer until signal appeared.

2.5.12 Generation of Unidirectional Deletions with Exonuclease III

DNA fragments of interest were cloned into pBluescript SK+. 10 μ l supercoiled plasmid DNA was double digested with restriction enzymes: one that generates a 4-base 3' protrusion (KpnI and SacI were used), and another which leaves a 5' protrusion adjacent to the insert from which deletions were to proceed. DNA was then phenol/chloroform extracted and ethanol precipitated.

Digested DNA was dissolved in 60 μ l 1 X ExoIII buffer and incubated on a 37°C water bath for at least 3 minutes. Then Exonuclease III was added to a concentration of 150 units per picomole recessed 3'termini, and rapidly mixed. 2.5 μ l samples were removed at 30 second intervals into individual tubes containing 7.5 μ l S1 nuclease mix on ice. After all samples were taken, these tubes were moved to a 30°C water bath and incubated for 30 minutes. The reaction was stopped by adding 1 μ l S1 stop buffer and heating at 70°C for 10 minutes. 2 μ l samples from each time point were analyzed on a 1% agarose gel. 1 μ l Klenow mix was added to each sample, and they were incubated for 10 minutes at 37°C. After adding 40 μ l ligase mix, samples were incubated at room temperature for 2 hours. Then 10 μ l of the ligation mixture of each time point was used to transform the E. coli strain NM522.

Solutions Used for Deletion:

Exo III buffer 10X:

660 mM Tris-Cl pH 8.0, 6.6 mM MgCl₂,

S1 buffer 7.4X:0.3 M K-acetate pH4.6, 2.5 M NaCl,

10 mM ZnSO₄, 50% (v/v) glycerol

S1 nuclease mix:

172 μ l dH₂O, 27 μ l S1 7.4X buffer, 60 units S1 nuclease

S1 stop buffer:

0.3 M Tris-base, 0.05 M EDTA

Klenow buffer:

20 mM Tris-Cl pH 8.0, 10 mM MgCl₂

Klenow mix:

30 μ l Klenow buffer,

3-5 units Klenow fragment of DNA polymerase I

dNTP mix:

0.125 mM each of dATP, dCTP, dGTP, and dTTP

Ligase 10X buffer:

500 mM Tris-Cl pH 7.6, 100 mM MgCl₂, 10mM ATP

Ligase mix:

790 μ l dH₂O, 100 μ l ligase 10X buffer,

100 μ l

1 50% (w/v) PEG6000, 10 μ l 100 mM DTT,

5 units T4 DNA ligase

2.5.13 DNA Sequencing by Dideoxynucleotide Chain Termination Method (Sanger, *et al.*, 1977)

a). Preparation of Template DNA

Since the phagemid pBluescript contains the origin of replication of the filamentous bacteriophage f1, single stranded DNA can be produced upon coinfection of a helper phage, for example, VCM13. Single colony containing desired phagemid was used to inoculate 5 ml TY-broth in a flask, and 40 μ l VCM13 helper phage (10¹⁰ pfu/ml) was added. After shaking at 37°C for 3 hours, Kanamycin was added to a concentration of 70 μ g/ml, and the culture was incubated overnight with vigorous agitation. Cells were pelleted by centrifugation at 12 Krpm for 5 minutes, and 1.2 ml supernatant was transferred into a fresh microcentrifuge tube. Phage particles were precipitated by adding 300 μ l 3.3 M amino-acetate, 20% (w/v) PGE6000, and incubated on ice for 15 minute, centrifuged for 15 minutes at 12 Krpm. The pellet was resuspended in 400 μ l TE and 0.1 volumes of Na-acetate (pH 5.2). DNA was phenol/chloroform extracted twice by vortexing each for 1 minute. DNA was ethanol precipitated, and dissolved in 25 μ l TE.

For double stranded templates, plasmid DNA was prepared by both mini-scale (2.5.6.1) and maxi-scale (2.5.6.2).

b). Sequencing reactions

For single stranded DNA, template and primer were annealed in a 14 μ l volume containing 8 μ l template DNA (1-2 μ g), 2 μ l annealing buffer, 2 μ l 17-oligonucleotide primer (0.2 mol/ μ l) at 65°C for 10 minutes and allowed to cool to room temperature slowly. For sequencing double stranded template, 5 μ l DNA (3-5 μ g) was mixed with 1 μ l 1N NaOH and 1 μ l primer (0.2 mol/ μ l), and incubated for 10 minutes at 37°C. The solution was neutralized with 1 μ l 1N HCl. Then 2 μ l annealing buffer was added.

To the annealed template and primer, 6 μ l "Enzyme Premix" was added and mixed gently. the reaction was incubated at room temperature for 5 minutes. 4.5 μ l (3.5 μ l for double stranded templates) was removed into individual tubes each containing 2.5 μ l T, C, G or A termination mix prewarmed to 38°C. The tubes were incubated at 37°C for 5 minutes, and the reaction was stopped by adding 5 μ l stop solution. The samples were heated at 75-90°C for 2-3 minutes, immediately before loading on to polyacrylamide gels.

c). Solutions for Sequencing

Annealing Buffer 5X:

200 mM Tris-Cl pH 7.5, 100 mM MgCl₂, 250 mM NaCl

Stop Solution:

95% Formamide (v/v), 20 mM EDTA,
0.05% (w/v) Bromophenol Blue,
0.05% (w/v) Xylene Cyanol FF

Labelling Mix 5X:

7.5 μ M dGTP, 7.5 μ M dCTP, 7.5 μ M dTTP

ddG Termination Mix:

80 μ M each dGTP, dATP, dCTP, dTTP; 8 μ M ddGTP

ddA Termination Mix:

80 μ M each dGTP, dATP, dCTP, dTTP; 8 μ M ddATP

ddT Termination Mix:

80 μ M each dGTP, dATP, dCTP, dTTP; 8 μ M ddTTP

ddC Termination Mix:

80 μ M each dGTP, dATP, dCTP, dTTP; 8 μ M ddCTP

Enzyme Premix:

1 μ l dH₂O, 1 μ l 0.1 M DTT, 2 μ l 1x labelling mix,
2 μ l T7 DNA polymerase (1.5 units/ μ l),
1 μ l (5 μ Ci) ³⁵SdATP

2.5.14 *In situ* Hybridisation to Third Instar Polytene Chromosome Preparations

2.5.14.1 Preparation of Third Instar Polytene Chromosomes

Microscope slides were washed in 1M HCl and then subbed by dipping into a solution of 0.1% (w/v) gelatin, 0.025% (w/v) chromic Potassium Sulphate, and air dried. Cover slips were washed in 1 M HCl, air dried and siliconised by dipping into a solution of Dimethyldichlorosilane (about 2% (v/v) in 1,1,1-trichloroethane) followed by air drying.

Salivary glands were dissected in Ringer's Solution (2.1) from late third instar larvae. The glands were then transferred to 20 μ l of 45% acetic acid on a microscope slide. A coverslip (18 mm²) was placed on top of the glands. The cells were broken by pressing the coverslip at one edge with a tissue followed by gentle tapping with a stiff, pointed probe starting over the glands and then gently working towards the edge in a spiral pattern. The slide was then turned over, onto a paper towel, and pressed at one edge so that the stain flowed to the opposite edge thus separating cell debris from the nuclei. The slide was inverted again, held at the edge with a paper towel and the chromosomes were spread by streaking across the coverslip in a zig-zag motion using a pointed probe. Finally the chromosomes were flattened by placing the slide between filter paper circles and pressing firmly. The slide was frozen in liquid nitrogen, the coverslip removed, and then the chromosomes were dehydrated in 95% (w/v) ethanol for 5 minutes, and air dried. The chromosomes were then examined under phase contrast microscopy and chosen for subsequent hybridisation experiments if the chromosome morphology was in good condition. The slides were stored at 4°C in an air tight dry plastic container.

2.5.14.2 Prehybridization of Polytene Chromosome Preparations

Before hybridisation, the slides were incubated in 2X SSC at 65°C for 30 minutes, then dehydrated through 2 changes of 70% (w/v) ethanol, 10 minutes each and 1 change of 95% (w/v) ethanol for 5 minutes and air dried. 200 μ l of 100 μ g/ml RNase A in 2 X SSC was pipetted onto the chromosomes and a large coverslip (25 mm²) was placed on top. The slides were incubated at 37°C in a moist chamber, then rinsed in 2 X SSC (3 changes at 5 minutes each) followed by dehydration through an ethanol series as described above. The chromosomal DNA was then denatured by boiling the slides in 10 mM Tris-Cl pH 8.0, 5 mM MgCl₂ for 2 minutes followed by dehydration as before.

2.5.14.3 Hybridization

The digoxigenin labelled RNA (2.5.8.3) was pelleted by centrifugation in a microfuge for 7 minutes, washed in 70% (w/v) ethanol, vacuum dried and resuspended in hybridisation buffer (50% (w/v) formamide, 4 X SSC). A volume of 30 μ l probe with approximately 2 μ g RNA was pipetted onto each squash, and covered with a plastic coverslip. The slides were sealed with cow gum (diluted with diethyl ether to a consistency easily pipetted using a glass pasteur pipette) and placed in a moist chamber at 42°C overnight.

2.5.14.4 Washing Steps After Hybridisation

After hybridisation the cow gum seal was easily removed using tweezers and the microscope slide floated off in the first wash. Slides were washed twice in 2X SSC, for 10 minutes, at room temperature. Slides were treated with RNase A (20 μ g/ml in 2 X SSC) at 37°C for 40 minutes, and subjected to following washes at room temperature: 2X5 minutes in 2X SSC, 2X5 minutes in PBS, 1X3 minutes in PBS with 0.1% (v/v) Triton-X-100, 3X5 minutes in PBS. Then slides were washed 2X5 minutes in buffer 1 (2.5.11.2), blocked for 30 minutes in buffer 2, washed 2X5 minutes again in

buffer 1. Slides were then submerged in antibody-conjugate (1:5000 in buffer 1) for 60 minutes. Unbound antibody-conjugate was removed by washing 2X15 minutes in buffer 1. Slides were incubated in buffer 3 for 2X15 minutes and covered with the colour detection solution (2.5.11.2) until signal appeared. Colour development was stopped in PBS.

2.5.14.5 Microscopy and Chromosome Staining

After developing, the slides were stained as follows. The slides were soaked in 10 mM Sodium Phosphate buffer, pH 7.0, for 3 minutes, then transferred to a 5% (v/v) dilution of Giemsa stain (supplied by SIGMA) in 10 mM Sodium Phosphate buffer pH 7.0. Slides were stained for 10 minutes, then placed with 10 mM Sodium Phosphate, pH 7.0 for 5 minutes and finally rinsed extensively in running water. The slides were air dried and viewed at X25, X40 and under immersion oil at X100 magnification to locate the precise band within which the signal was present. The in situ hybridisation signals were photographed.

2.5.15 *In situ* Hybridization for Localization Specific RNA in *Drosophila* Ovaries, Testes and Embryos

2.5.15.1 Tissue Preparation

Adult flies were fed on cornmeal food with supplement of dry yeast for 4-5 days after eclosion. Ovaries and testis are hand-dissected in Ringers' solution.

Eggs were collected in a fly cage bottomed with a Petri dish containing Adh fly food with yeast paste spreaded on the middle in a period of 16 hours at 25°C. The eggs were washed with water and dechorioned in 5% sodium hypochlorite for 5 minutes. the embryos then washed with 0.1% Triton X-100.

2.5.15.2 Fixation

The ovaries, testes and embryos were fixed in the same way: the tissues (i.e. ovaries , testes and embryos) were transferred into glass scintillation vials containing 1.6 ml of 0.1 M Hepes, pH 6.9, 2 mM MgSO₄, 1 mM EGTA. To this vial 0.4 ml freshly made 20%

(w/v) paraformaldehyde and 8 ml heptane was added. The vial is shaken vigorously for 15-20 minutes to maintain an effective emulsion. Then most of the solution was discarded, and 10 ml methanol was added. After another change of fresh methanol, methanol was replaced with 10 ml 1:9 DMSO/methanol, and the tissues were stored at -70°C for 1 hour.

2.5.15.3 Pretreatment

The tissues were transferred into an 1.5 ml microcentrifuge tube. They were first washed 3 times each for 5 minutes in 1 ml PBT(PBS+0.1% (v/v) Tween 20). Then they were treated with 50 µg/ml Proteinase K in PBS. For different tissues the time of Proteinase K digestion was varied: 30 minutes for ovaries, 3-5 minutes for testes and embryos. The proteinase K was inactivated by incubating for 2 minutes in 2 mg/ml glycine in PBT and the tissues were washed 2 times in PBT, refixed for 20 minutes in PP(4% paraformaldehyde in PBS) and washed again 3 times each for ten minutes in PBT.

2.5.15.4 Hybridization and Washing

The hybridization solution (HS) consists of 50% (v/v) formamide, 5 X SSC, 50 µg/ml heparin, 0.1% (v/v) Tween 20 and 100 µg/ml sonicated and denatured salmon sperm DNA. The tissues were washed for 20 minutes in 1/1 HS/PBT and then 60 minutes in HS. The tissues were prehybridized in HS at 45°C for 40 minutes. After prehybridization most HS was removed and denatured probe was added and mixed thoroughly. The probe concentration was about 0.5 µg/ml in HS. The probe was heat denatured(95°C for 10 minutes then stands on ice for 3 minutes) with the presence of 10 µg sonicated salmon sperm DNA. Hybridization was conducted over night at 45°C in a water bath. The tissues were then washed at 37°C for 20 minute each for following steps: 20 minutes in HS, 20 minutes in 4/1 HS/PBT, 20 minutes in 3/2 HS/PBT, 20 minutes in 2/3 HS/PBT, 20 minutes in 1/4 HS/PBT, 2 times each for 20 minutes in PBT.

2.5.15.5 Signal Detection

The tissues were incubated for 1 hour at room temperature with gentle shaking in 500 μ l antibody-conjugate solution (purchased from Boehringer and freshly diluted to 1/5000 in PBT). Then the tissues were washed 4 times each in PBT for 20 minutes, 3 times for 5 minutes each in 100 mM NaCl, 50 mM MgCl₂, 100 mM Tris-HCL pH9.5, 1m M Levamisole, 0.1% Tween 20. The tissues were then transferred into a small beaker containing 1 ml of above solution and 4.5 μ l NBT and 3.5 μ l X-phosphate solutions were added and mixed thoroughly. The signal was developed in dark for 10-30 minutes and stopped in PBT. The tissues were either directly examined and photographed under a microscope in PBT, or mounted in water mount media and also photographed using colour film (Kodak Ektachrome 160).

CHAPTER THREE
***IN SITU* DETECTION OF ENHANCERS ACTING IN ADULT**
GONADS OF *DROSOPHILA MELANOGASTER*



3.1 General Description of the *LacZ* Staining Patterns

Adult males and females of 61 strains carrying an insertion of the enhancer detector P[*lac*, *ry*⁺]A (O'Kane and Gehring, 1987; also see the Introduction) were stained with X-gal, to examine the β -galactosidase expression pattern with the aim of identifying genes which are specifically active in the adult germ-line cells. These strains were chosen from 230 transposant lines which were previously stained for β -galactosidase activity, and show gonad specific expression of β -galactosidase (Bownes, 1990). The majority of these strains (56 out of 61) show the *lacZ* staining in either male or female gonads, or both. In most of the strains showing positive staining, the β -galactosidase expression is either temporally or spatially controlled. There were two exceptions in which a high level of β -galactosidase expression is detected in all cells of both male and female gonads. A common feature of the staining pattern is that in most strains the *lacZ* fusion protein is localized in the nucleus of the stained cells. It is proposed that the *lacZ* fusion protein contains a nuclear targeting sequence derived from the P-transposase protein (Grossniklaus *et al.*, 1989). Table 3.1 is a list of the staining data for the 61 transposant strains. For convenience in describing the strains, the *lacZ* staining patterns in the ovary and male gonad are described separately.

3.2 *lacZ* Staining Patterns of the Ovary

Based on the cell types, the patterns of β -galactosidase expression in these strains can be broadly divided into three groups: somatic specific, germ-line specific and those showing the *lacZ* staining in both somatic cells and germ-line cells. In some strains, the *lacZ* fusion gene is expressed in the germarium and in other cells in the later stages of oogenesis. In considering the lineage between the cells in the germarium and their descendants in later stages, these strains are classified according to their staining at the later stages. Table 3.2 shows the distribution of the *lacZ* staining in different cell types.

3.2.1 Somatic Specific Staining Patterns

Table 3.1 A list of the lacZstaining patterns in 61 transposant strains.

| <u>Number</u> | <u>line</u> | <u>female</u> | <u>male</u> |
|----------------------|--------------------|---|---|
| 1. | A18 | - | testis epith. |
| 2. | B2 | S10-11 nurse cell nucl. S12-14 ooplasm | cysts, sperm bundles |
| 3. | B12 | s1-6 polar cells | tip of testis, testis epith., ED, TD |
| 4. | B19 | - | - |
| 5. | B21 | - | testis epith. |
| 6. | B19 | - | - |
| 7. | B23 | S10-13 2 border cells | testis epith. |
| 8. | B27 | S10 Nurse cell nuclei | - |
| 9. | B31 | Germarium, S12-13 ooplasm | cyst cells, tip of ACG, TD |
| 10. | B32 | - | - |
| 11. | C2 | Germarium, border cells | tip of testis, TD |
| 12. | C3 | - | sperm bundles |
| 13. | C4 | - | - |
| 14. | C5 | S10ab follicle cells | - |
| 15. | C7 | - | testis epith. |
| 16. | C20 | germarium, S10 n/o junction | tip of testis |
| 17. | C25 | S10ab 5-6 border cells | testis, epith., TD, SV. |
| 18. | C26 | S10ab follicle cells | - |
| 19. | C27 | - | tip of ACG |

Table 3.1 continued

| <u>Number</u> | <u>line</u> | <u>female</u> | <u>male</u> |
|---------------|-------------|---|-----------------------------|
| 20. | C28 | all cell in ovary | all cells in male gonad |
| 21. | C29 | germarium, S10-12 nurse nuclei, S12-14 ooplasm | spermatocyte, SV, AED, EB |
| 22. | C32 | nurse cell nuclei, follicle cells | testis epith. |
| 23. | C36 | S10 nurse cell nuclei | testis epith. |
| 24. | C37 | S10 posterior polar cells, follicle cells at n/o junction | - |
| 26. | C38 | S14 ooplasm | tip of ACG |
| 27. | C42 | oviducts | AED |
| 28. | C55 | germarium, S10 follicle cells | tip of testis, cyst cells |
| 29. | D6 | follicle cells in all stages | - |
| 30. | D7 | - | - |
| 31. | D8 | germarium, S10 follicle at n/o junction | TD, AED |
| 32. | D13 | S4-12 follicle cells | testis epith. |
| 33. | D14 | oocyte, nurse cells, follicle cells | - |
| 34. | D16 | - | spermatocyte |
| 35. | D19 | germarium, S7-11 nurse nuclei, oocyte nuclei | spermatocyte, sperm bundles |
| 36. | D22 | germarium | - |

Table 3.1 continued

| <u>Number</u> | <u>line</u> | <u>female</u> | <u>male</u> |
|---------------|-------------|---|--|
| 37. | D83 | germarium, S9-11 nurse nuclei, oocyte nuclei | testis epith. |
| 38. | E10 | S9-13 nurse cell nuclei S11-14 ooplasm S10 follicle cells | all cells in male gonad |
| 39. | F1 | S1-4 polar cells | - |
| 40. | F2 | S10-11 squamous follicle cells, S10-13 columnar follicle cells | - |
| 41. | F9 | S1-6 polar cells S9-11 squamous and columnar follicle cells, follicle cells at n/o junctionn | testis epith. |
| 42. | F13 | germarium, S1-7 polar cells S8-14 columnar follicle cells | cyst cells |
| 43. | F14 | S1-11 nurse cell nuclei, oocyte nuclei, S12-14 ooplasm | ACG |
| 44. | F16 | S10 nurse cell nuclei | testis epith. |
| 45. | F17 | nurse cells, follicle cells | - |
| 46. | F21 | nurse cells, oocyte, follicle cells | testis epith. |
| 47. | F22 | - | spermatocyte |
| 48. | F24 | S10 follicle cells | testis epith., tip of testis, tip of ACG |
| 49. | G16 | S9-11 nurse cell nuclei, S10-12 oocyte nuclei, S11-14 ooplasm | - |

| <u>Number</u> | <u>line</u> | <u>female</u> | <u>male</u> |
|---------------|-------------|--|------------------------------|
| 50. | G18 | - | - |
| 51. | H5 | S1-14 follicle cells | testis epith. |
| 52. | H10 | S13-14 follicle cells | testis epith. |
| 53. | H12 | S9-11 nurse cell nuclei S10-11 oocyte nuclei, s12-14 ooplasm | spermatocyte |
| 54. | H14 | S10-13 border cells, posterior polar cells | testis epith. SV, AE, ACG |
| 55. | J1 | S1-7 follicle cells | testis epith. |
| 56. | J2 | - | spermatocyte |
| 57. | J3 | S10-11 nurse cell nuclei S12-14 ooplasm | Spermatocyte |
| 58. | J5 | S9-11 follicle cells, S13 appendage cells | testis epith. |
| 59. | O | - | testis epith, SV |
| 60. | Q | S11-13 border cells | testis epith, SV |
| 61. | T | S13 appendage cells | - |

ACG= accessory gland

AE= ejaculatory duct

epith = epithilum

o/n junction = oocyte/nurse cell junction

SV = seminal vesicle

Twenty three strains carrying P[*lac*, *ry*⁺]A show the β -galactosidase activity in the ovarian follicle cells. Although the size of the sample is small, some interesting subsets of the follicle cells have been revealed by the staining. There are 6 (B23, C2, C25, D8, H14, Q) strains that show the β -galactosidase activity in the border cells, or the follicle cells situated at the oocyte-nurse cell junction (Figure 3.1). Among the strains showing the *lacZ* activity in the border cells, different numbers of the border cells expressing P[*lac*, *ry*⁺] were stained in different strains, ranging from 2 to 10 border cells. Temporal variations were also observed, for example, in one strain (C25) the border cells were stained in stage 10 egg chambers, while the *lacZ* fusion gene was expressed in stage 13 in another strain (Q). Grossniklaus *et al.* (1989) reported that a large proportion of their transposants expressed β -galactosidase in their border cells. As the functions of these follicle cells are more complex than other follicle cells, the high frequency and diversity of staining in these cells may simply reflect the complexity of gene activity in these cells, and different functions between these follicle cells. Five strains (B12, C37, F1, F9, F13) have β -galactosidase activity in the follicle cells at both anterior and posterior end of the egg chambers during stage 1-6 of oogenesis (Figure 3.2A). These are called polar cells. Expression was found in early stages of oogenesis (stage 1-6). Two of these strains also showed the *lacZ* staining in later stages. One (F9) had its squamous follicle cells and the cells at the oocyte-nurse cell junction stained during stage 9-11 (Figure 3.2B), while the other (C37) expressed the *lacZ* fusion gene in the columnar follicle cells during stage 9-14 (Figure 3.2C)

There is one strain (H14) that expresses β -galactosidase in the follicle cells at both the anterior and posterior ends of the oocyte from stage 8 to stage 13 (Figure 3.3). This expression pattern is interesting because activation of some maternal-effect genes in the embryo requires certain pre-positioned gene products, which act upstream of these genes. In the case of a terminal group maternal gene, the torsal protein is activated at the both anterior and posterior ends of the embryo by the product of the torsal-like gene (Sprenger *et al.*, 1989). The torsal-like product is synthesised in the

Table 3 2.

| Cell Type Stained | Strain | No |
|---|------------------------|----|
| Germarium | D22 | 1 |
| germarium/ oocyte | B31 | 1 |
| Germarium/ nurse cells | | |
| Germarium/ follicle cells | C2, C20, D8 | 3 |
| Germarium/ oocyte/ nurse cells | D19, D83 | 2 |
| Germarium/ oocyte/ follicle cells | C29 | 1 |
| Germarium/ oocyte/ nurse cells / follicle cells | C28 | 1 |
| oocyte | C38 | 1 |
| Nurse cells | B27, C36, F16 | 3 |
| follicle cells | | |
| oocyte/ nures cells | B21, F14, G16, H12, J3 | 3 |
| oocyte/ follicle cells | | |
| oocyte/ nurse cells / follicle cells | D14, E10, F21 | 3 |
| Nurse cells/ follicle cells | C32, F17, J1 | 3 |
| Oviducts | C42 | 1 |

Figure 3.1

LacZ staining of the transposant strains

A. Strain Q

The *lacZ* fusion protein is detected in the border cells (arrow).

Scale bar = 50 μm

B. Strain C20

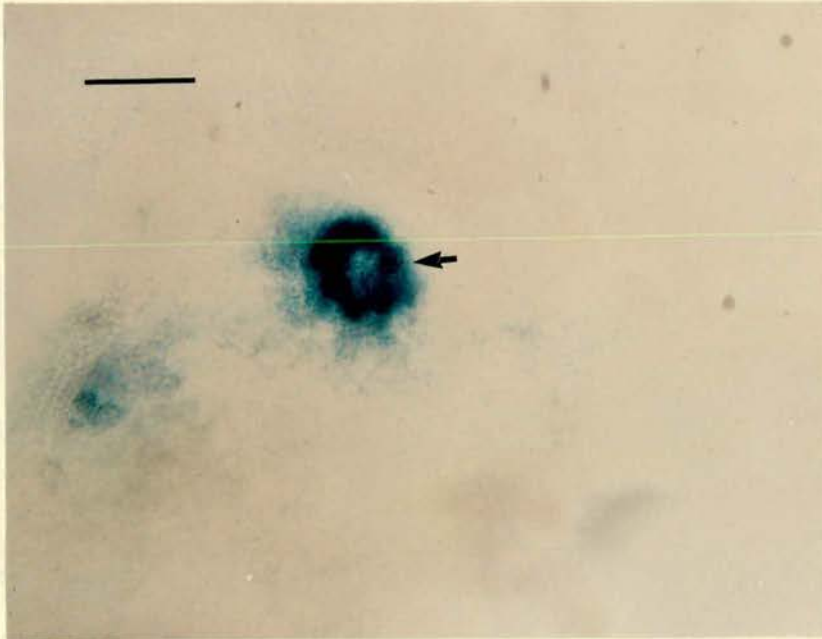
The *lacZ* fusion protein is detected in the cells at the oocyte-nurse celljunction. Scale bar = 50 μm

O = oocyte

N = nurse cells

FIGURE 3.1

A



B



Figure 3.2

lacZ staining of the transposant strains

A. Strain F9

The *lacZ* fusion protein is detected in stage 1-6 polar cells (arrows). Scale bar = 50 μ m

B. Strain F9

The *lacZ* fusion protein is detected in the squamous follicle cells in stage 10 egg chamber. Scale bar = 50 μ m

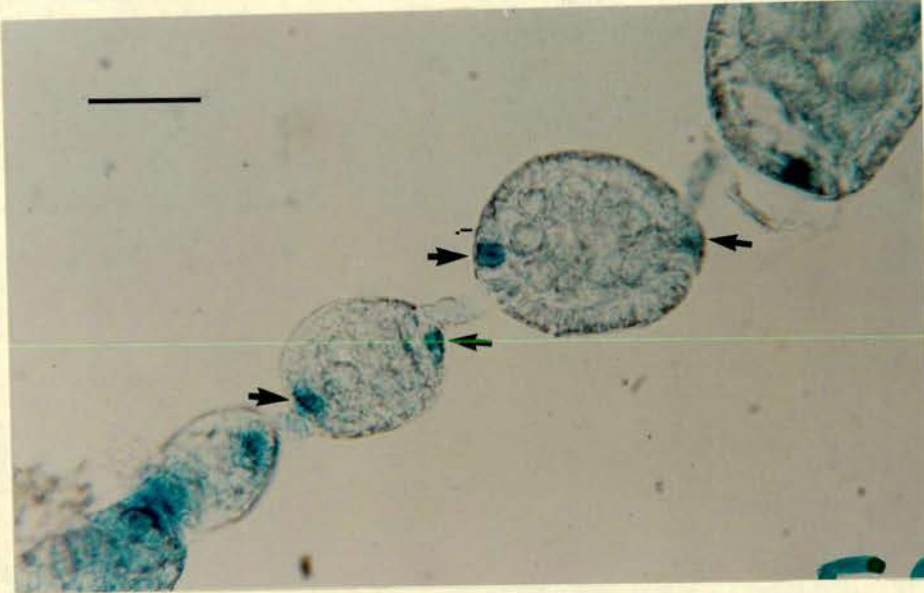
O = oocyte

N = nurse cells

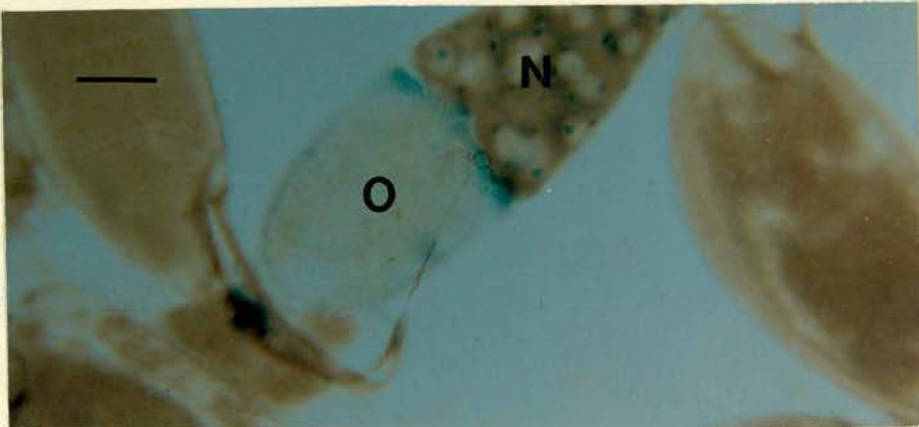
C. Strain C37

The *lacZ* fusion protein is detected in the nuclei columnar follicle cells in stage 14 egg chamber. Scale bar = 50 μ m

FIGURE 3.2



A



B



C

Figure 3.3

lacZ staining of the transposant strain H14

The *lacZ* fusion protein is detected in the border cells and the follicle cells (arrows) at the posterior end of the oocyte.

Scale bar = 50 μ m

O = oocyte

N = nurse cells

FIGURE 3.3



Figure 3.4

lacZ staining of the transposant strains

A. Strain F24

The *lacZ* fusion protein is detected in the columnar follicle cells surrounding the oocyte in stage 10. Scale bar = 50 μm

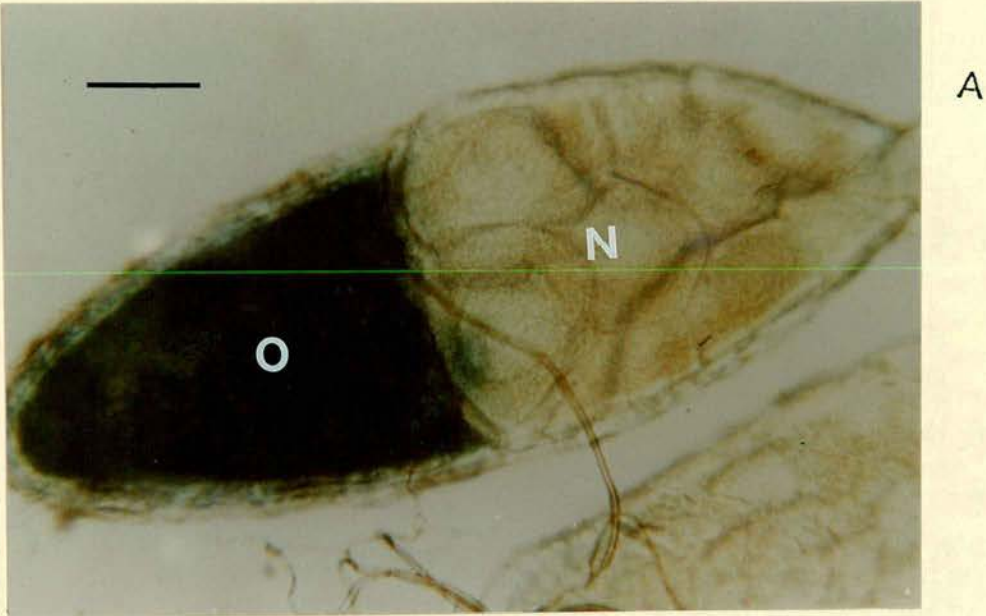
O = oocyte

N = nurse cells

B. Strain J5

The *lacZ* fusion protein is detected in the cells surrounding the dorsal appendages (arrows). Scale bar = 50 μm

FIGURE 3.4



follicle cells at both anterior and posterior ends of the oocyte, and pre-positioned in the mature egg during oogenesis (Stevens *et al.*, 1990). The expression of the *lacZ* fusion gene in such a way reflects the activity of a certain gene which functions in a way similar to that of torsal-like gene. These follicle cells may also play an important role in determining the overall polarity of the embryo along the anterior/posterior axis.

In the remaining 9 strains (C5, C26, C55, D6, D13, F24, H5, H10, J5) the *lacZ* fusion protein is detected in the columnar follicle cells surrounding the oocyte. The expression patterns in these strains vary at different stages during oogenesis. One strain (H5) expressed the *lacZ* fusion gene from stage 4 to 12. Three strains showed *lacZ* staining during stage 10-11 (Figure 3.4A). The *lacZ* staining was also observed in later stages. In one strain (J5), the fusion gene was expressed in the columnar cells during stage 9-10, and later during stage 13-14, it was expressed in the cells surrounding the dorsal appendages (Figure 3.4B)

3.2.2 Germ-Line Specific Staining Patterns

Thirteen strains express β -galactosidase in the germ-line cells. The expression patterns also vary between different cells and different stages.

Of the thirteen strains which show germ-line specific staining, nine (B2, B27, C29, C36, D19, D33, F16, G16, J3) express the *lacZ* fusion protein in the nurse cells during stage 7-11 of oogenesis (Figure 3.5A). Grossniklaus *et al.* (1989), and Fasano and Kerridge (1988) reported that large numbers of transposants express β -galactosidase in the nurse cells during these stages. This is presumably due to the fact that nurse cells are the most active in transcription during these stages.

In some strains, β -galactosidase is detected in the oocyte nucleus but not later than stage 11 (Figure 3.5A). In all the strains that express β -galactosidase in the oocyte nucleus, the *lacZ* activity is also detected in the nurse cells. It may be that β -galactosidase is synthesised in the nurse cells and later translocated in the oocyte nucleus, since the data from autoradiography has shown that there

Figure 3.5

lacZ staining of the transposant strains

A. Strain D19 (ovary)

The *lacZ* fusion protein is detected in nurse cell and oocyte nuclei during stage 7-11. Weak expression can also be seen in the germarium. Scale bar = 50 μm

O = oocyte

N = nurse cells

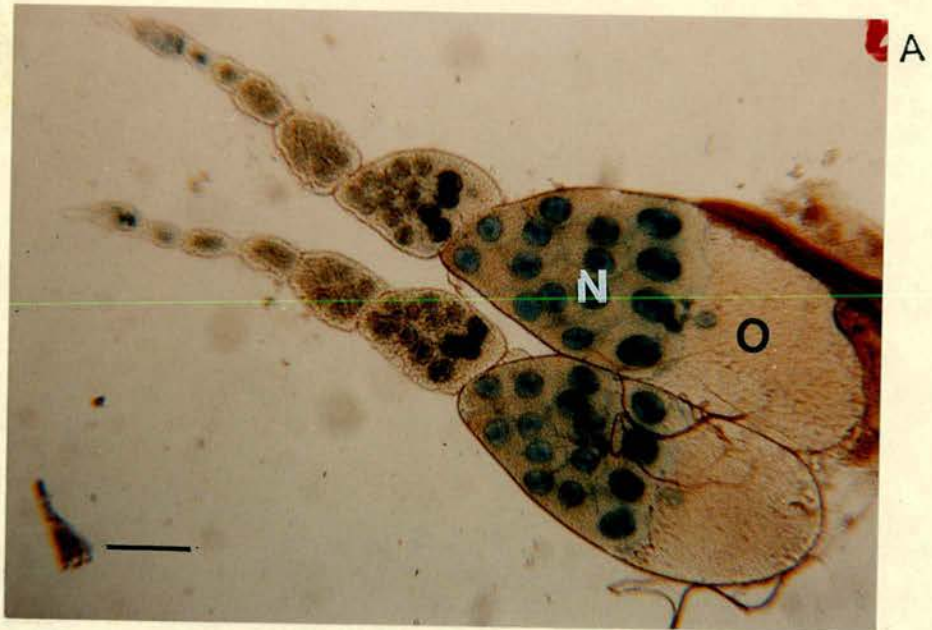
B. Strain D19 (testis)

The *lacZ* fusion protein is detected in the spermatocyte cysts, and sperm bundles. Scale bar = 50 μm

SC = spermatocyte cyst

SB = sperm bundle

FIGURE 3.5



is little transcriptional activity in the oocyte nucleus during oogenesis (Mahowald and Tiefert, 1970). Alternatively transcription in oocyte nucleus may have previously been underestimated. Grossniklaus *et al.* (1989) suggest that the presence of β -galactosidase in the oocyte nucleus reflects the gene activity of the oocyte nucleus itself. Evidence supporting this interpretation comes from the observation that some of their transposants with a high level of expression of β -galactosidase in the nurse cells do not show the *lacZ* staining in the oocyte. In our analysis, 3 strains (B27, C36, F16) were observed which strongly express β -galactosidase in the nurse cells but not the oocyte (data not shown). Thus we also favour the interpretation that the activity of the oocyte nucleus has been previously underestimated.

3.3 *lacZ* Staining Patterns of Male Gonads

In the male gonad, β -galactosidase is expressed in different compartments, and different types of cells within certain compartments. Three transposant strains (C2, C20, C55) express β -galactosidase at the anterior tip of the testis (Figure 3.6B). Yet it is difficult to determine whether these cells are germ-line cells or somatic cells. The expression of the *lacZ* fusion gene is found in later stages of spermatogenesis in these strains. Yet none of them show staining in the germ-line cells. In the 23 strains expressing the *lacZ* fusion gene in the somatic cells of testis, the majority of them show the *lacZ* staining in the epithelium cells (Figure 3.7A). The staining varies between different strains in terms of intensity and location. Six strains (B31, C27, C38, F14, F24, H14) show the staining in the accessory glands (Figure 3.7B). Three of them (B31, F24, H14) also express β -galactosidase in the cyst cells and testis epithelium. There are nine strains (B2, C3, C29, D16, D19, F22, H12, J2, J3) expressing β -galactosidase in the germ-line cells of adult testes. Seven of them show the *lacZ* staining in the spermatocyte cysts (Figure 3.8A), two of them also express the *lacZ* fusion gene in the sperm bundles (Figure 3.5B). A higher number of the transposants express the *lacZ* fusion gene in the spermatocyte cysts, a fact agreeing with previous observations that most transcription ceases

Figure 3.6

lacZ staining of the transposant strains

A. Strain C55 (ovary)

The *lacZ* fusion protein is detected in the germarium (arrow).

Scale bar = 50 μm

B. Strain C55 (testis)

The *lacZ* fusion protein is detected in the tip of testis (arrow).

Weak expression can be seen in the cyst cells. Scale bar = 50 μm

FIGURE 3.6

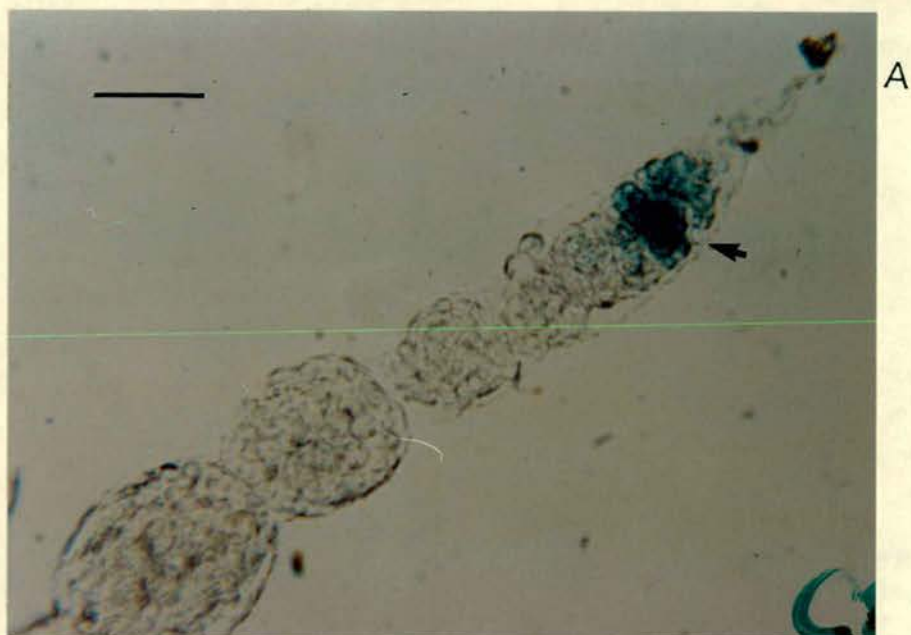


Figure 3.7

***lacZ* staining of the transposant strains**

A. Strain F24

The *lacZ* fusion protein is detected in the epithelium of the testis (blue spots). Scale bar = 50 μ m

B. Strain C27

The *lacZ* fusion protein is detected in the tips of the male accessory glands (AG). Scale bar = 50 μ m

FIGURE 3.7

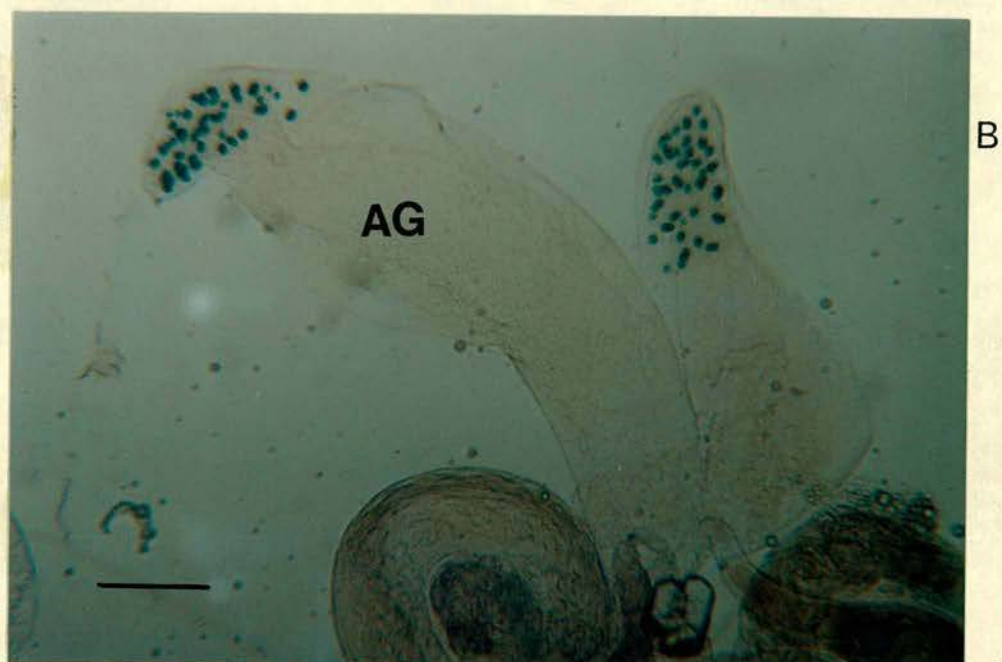
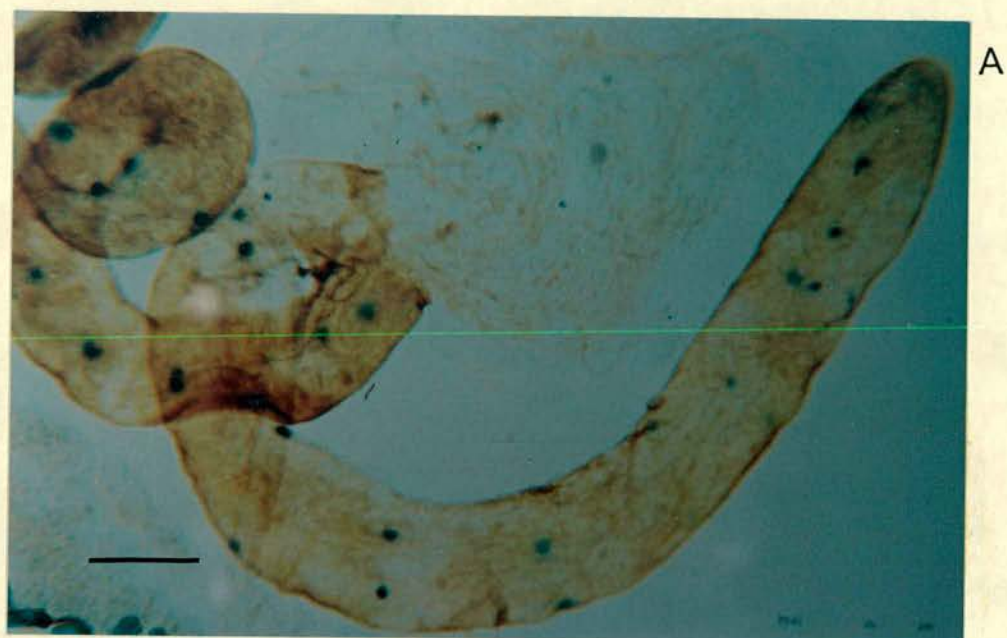


Figure 3.8

lacZ staining of the transposant strains

Strain F22

The *lacZ* fusion protein is detected in the spermatocyte cysts of the testis (arrows). Scale bar = 50 μ m

FIGURE 3.8



before meiosis (Henning, 1967; Brink, 1968). However, in the case of C3 and D19 (figure 3.5B), the expression of the *lacZ* fusion in sperm bundles suggests that there may still be transcriptional activity during the post-meiotic stages of spermatogenesis. Similar staining patterns have been observed in other strains in this laboratory (Bownes, 1990)

3.4 Comparison of the Expression Patterns of β -galactosidase between Male and Female Gonads

A comparison of the β -galactosidase expression between male and female gonads reveals several facts. A large proportion of the transposant strains (38 out of 56) express the *lacZ* gene in the gonads of both sexes. The number of strains showing sex-specific staining is relatively small; nine are male-specific, and ten are female-specific. This data can be interpreted in two ways: firstly, the large number of non-sex-specific staining may indicate that there are fewer genes functioning in a sex-specific manner. Secondly, the expression of the *lacZ* fusion gene in male and female gonads may be under the control of different enhancers in each sex, since enhancers can act over a distance and in different orientations (Jeang K.-T. and Khouty, 1988), ie they are showing the expression pattern of more than one gene.

Among the 23 strains which express β -galactosidase in the ovarian follicle cells, 20 strains also express the same enzyme in the testes epithelial cells or other somatic cells of the male gonad. But none of these strains express β -galactosidase in the germ-line derived spermatocyte cysts. This implies that there are some common lineages between male and female somatic gonadal cells. This is clearly shown by the *lacZ* staining pattern in strain C42 (Figure 3.9AB); the *lacZ* fusion protein labelled oviducts in female and anterior ejaculatory duct in male. These two organs are all derived from larval genital discs (Babcok, 1971ab; Lindsley and Tokuyasu, 1980). In the ten strains (B2, B27, C25, C29, C36, D19, D33, F14, F16, G16, J3) showing the *lacZ* staining in female germ-line cells, four them (B2, C29,D19, J3) express β -galactosidase in spermatocyte cysts. These results suggest that there are some

Figure 3.9

***lacZ* staining of the transposant strains**

Strain C42

A. The *lacZ* fusion protein is detected in the lateral and common oviducts (arrow). Scale bar = 100 μm

B. The *lacZ* fusion protein is detected in the ejaculatory duct in male gonad (arrow). Scale bar = 100 μm

FIGURE 3.9



A



B

enhancers which are specifically active in germ-line cells of both sexes. Some maternal-effect genes are known to be transcribed in a similar fashion. For example, the gene *vasa* is expressed in both male and female germ-line cells, yet its activity is only essential in the female germ line, some mutants are female sterile not male (Hay *et al.*, 1988; Lasko and Ashburner, 1990). There may be some common factors existing in both male and female germ-lines in order to confer the enhancer activities in the germ-line cells of both sexes. But the product of a gene activated by the enhancer may only be required in one sex even though it is active in both sexes.

In the strains showing male specific expression of β -galactosidase, four strains (C3, D16, F22 and J2) stained in spermatocyte cysts, and one of them (C3) also expressed β -galactosidase in the sperm bundles. The five somatic specific strains expressed the *lacZ* fusion gene in different cells, one (C27) is in the cells at the tip of the accessory glands (Figure 3.7B), three are in the cyst cells and one is in the testis epithelial cells (Figure 3.7a). Of the ten female specific strains, six show (C5, C37, D6, F1, F2, T) the *lacZ* staining in the follicle cells, two (B27, G16) are germ-line specific, and two (D14, F17) express the *lacZ* fusion gene in both follicle cells and nurse cells. The higher number of strains showing *lacZ* staining in the follicle cells

3.5 The Expression Pattern of β -Galactosidase in Strain D19

In this section, I present the result of the *lacZ* staining of one transposant strain D19 which was chosen for further molecular analysis.

The *lacZ* fusion protein is detected in the ovary and testis in this strain (Figure 3.5a.b). In the ovary, there is a weak expression of β -galactosidase in the germarium. Starting from stage 7 of oogenesis, the *lacZ* fusion protein was detected in the nuclei of nurse cells. The staining intensity increased with the growth of the egg chamber. The expression of β -galactosidase peaked at stage 10 to stage 11 of oogenesis. The *lacZ* fusion gene is expressed most strongly in the nurse cells near the oocyte. This is obvious within the stage 10 or

stage 11 egg chambers. β -galactosidase is also detected in the oocyte nucleus in these stages. No *lacZ* staining is found in egg chambers older than stage 11. As described in the Introduction, RNA transcription in the nurse cells increases from stage 7 onwards and reaches maximum at stage 11, and most of these genes are acting maternally to support the growth of the growing oocyte and provide materials for embryogenesis. The expression of the *lacZ* fusion D19 is consistent with general pattern of most genes active in the nurse cells. The detected enhancer probably acts on certain genes, with the maternal effect for the oocyte or the embryo. In testes, the *lacZ* fusion gene is expressed in the spermatocyte cysts and sperm bundles. The *lacZ* fusion protein has not been detected in other tissues, so the expression of the *lacZ* fusion gene is germ-line-specific

3.6 Discussion

The results of staining 61 transposant strains, carrying the enhancer detector P[lac, ry⁺], for the β -galactosidase activity in the adult gonads are described. Analysis of the staining patterns has provided important information about the gene regulation in ovaries and testes. Although the sample size is relatively small, as the primary concern of this project is not a general survey of the staining pattern, very specific and informative staining patterns in respect to the gene regulation in male and female gonads have been identified.

In the ovary, 23 of 47 strains express β -galactosidase in the follicle cells. In 13 strains the enzyme activity is detected in germ-line cells, two strains show the *lacZ* staining in all cells of the ovary.

Some strains show a unique temporal and spatial pattern of the *lacZ* staining in follicle cells. The *lacZ* fusion protein labels different groups of follicle cells. The enzyme activity is detected in the columnar cells (Fig.3.4A); the squamous cells (Fig.3.2B); the polar cells (Fig.3.2A); the cells of dorsal appendages; the border cells (Fig.3.1A) and the cells at the oocyte-nurse cell junction (Fig.3.1A). Given that the enhancer trap technique reveals the existence of different enhancer elements (O'Kane and Gehring, 1987), the

different staining patterns generated by the enhancer detector P[lac, ry⁺] must represent different gene activities in the follicle cells. As described in the introduction, different groups of follicle cells play different roles in specifying the complex structure of the egg shell. Except for the function involving the construction of the egg shell, the follicle cells may possess important functions in determining the polarity and pattern formation of the embryo by specifying the activity of some maternal effect genes and their products. There are several strains, in which the *lacZ* fusion gene is active in the polar cells and follicle cells at the anterior or posterior poles of the oocyte (Figure 3.2C). Brower *et al.* (1981) reported the spatial location of an antigen which labels predominantly the polar cells. There is clear evidence that follicle cells at the anterior and posterior poles of the oocyte are involved in the expression of the maternal gene *torsal-like*. The product of *torsal-like* acts as a ligand to activate *torsal*, another maternal effect gene required for the determining the terminal structure of the embryo (Stevens, 1990). In different strains, the *lacZ* fusion protein marks different polar cells, and this may reflect the inherent polarity of the follicle cells with respect to gene activity.

In the germ-line, the *lacZ* staining pattern is relatively simple. The fusion protein is found in germarium, nurse cells and oocyte. Nonetheless, previous genetic and molecular data suggest that there are large numbers of genes active in the nurse cells during oogenesis. Evidence from genetic studies suggests that 50% (Perrimon and Gans, 1983) to 75% (Perrimon *et al.*, 1986) of the recessive female-sterile mutations and 70% of the zygotic lethal mutations on the X-chromosome affect the gene expressed in germ-line cells. In the survey of β -galactosidase expression in 600 transposants carrying the enhancer detector P[lArB], Grossniklaus *et al.* (1989) have found that the transposants which express the *lacZ* fusion gene in the germarium and nurse cells comprise 76% of the transposants which express the fusion protein in ovarian cells. Compared with the follicle cells, germ-line cells in the ovary differentiate a smaller range of cell types; the stem cell only produces one oocyte and 15 nurse cells. This may explain why the

staining pattern in the germ-line is simpler than that in the follicle cells.

As shown in table 3.1, in a few strains β -galactosidase labels different groups of cells in the ovary and testis. In some strains, the enzyme marks the germarium and follicle cells, or nurse cells at the later stages of oogenesis. These staining patterns may be an indication of the lineage between the cells in the germarium and their descendants. Another example is the strain which shows *lacZ* staining in polar cells in stage 1-6 of oogenesis, and the *lacZ* fusion protein which marks the squamous follicle cells and the follicle cells at the oocyte-nurse cell junction. Some follicle cells at the anterior end of the egg chamber migrate in such a way (Margaritis *et al.*, 1980).

Interestingly, in some strains the *lacZ* fusion gene is expressed in the cells with an obviously different origin. In the strain F14, the *lacZ* fusion protein is detected in the male accessory gland and in female germ-line cells (data not shown). There are three possible explanations for this staining pattern. Firstly, the *lacZ* fusion gene is under the control of different enhancer elements which control the expression of different genes in different tissues, and it happens that the P[lac, ry⁺] transposon is just inserted at a position where it is exposed to the different enhancers. Secondly, it is possible that one gene has more than one enhancer which confers expression in different tissues. This is true for the *yp1* and *yp2* genes as described in the Introduction. These two genes share two enhancers; one specifies their fat body expression while the other confers transcription in the ovarian follicle cells (Carabedian *et al.*, 1985, 1986). Finally, this expression pattern could be a result of multiple insertions of P[lac, r⁺].

CHAPTER FOUR

CLONING OF THE DNA SEQUENCE FLANKING THE P[LAC, RY⁺]A INSERTION IN STRAIN D19

4.1 Introduction

As described previously, some lines with very specific and interesting staining patterns were found. The line D19 was chosen for the following reasons: 1. the β -galactosidase expression is germ-line specific; 2. the expression of the *LacZ* fusion gene is very strong, so abundant transcripts could be expected.

Many approaches could be taken to clone the genomic sequences flanking the transposon either at the 5' end or the 3' end. These may contain genes showing the same expression pattern as the *lacZ* fusion gene. The direct way is to use the EcoRI or HindIII restriction sites within the transposon and clone the upstream genomic region flanking the transposon. A genomic library can be constructed and then the nearby tagged DNA is isolated by either by using the P-element or the *lacZ* gene as a probe.

4.2 Construction and Screening of the D19 Genomic Library

In order to choose an appropriate vector for cloning, DNA samples from flies of the D19 and wild-type, Oregon R strains were digested with EcoRI and electrophoresed in an agarose gel, and transferred onto a Hybond-N nylon membrane (Amersham) in a Southern blot. The filter was probed with a radio active probe made from a 584bp HindIII fragment containing the 5'-end sequence of the P-element (not including the terminal repeats) from the plasmid Carnegie 20 (Rubin and Spradling, 1983). Figure 4.1A shows the results. In lane 2 there are no positive signals detected from the Oregon R flies, whereas in lane 1 the P-element probe hybridized to a 2.1kb EcoRI fragment of DNA from D19 flies. This 2.1 kb fragment contains the 5'region of P-element sequence and the flanking genomic DNA.

The vector λ NM1149 (Murray, 1983) was chosen for construction of the genomic library from D19. This was based on information obtained from the Southern analysis described above. The vector has the cloning capacity of DNA inserts of 0-15 kb. The unique EcoRI and HindIII cloning sites lie within the lambda cI gene. Only when this gene is disrupted can the phage form plaques

Figure 4.1

A. Southern blot of genomic DNA digested with EcoRI, probed with the

584 bp 5'end of P-element in plasmid Carnegie 20

1 = D19

2 = Oregon R

B. Ethidium Bromide stained agarose gel containing EcoRI, and EcoRI/HindIII digested λ D19RI1. HindIII digested Carnegie 20

1 = 1 kb ladder marker

2 = λ D19RI1, EcoRI digestion

3 = λ D19RI1, EcoRI/HindIII digestion

4 = Carnegie 20, HindIII digestion

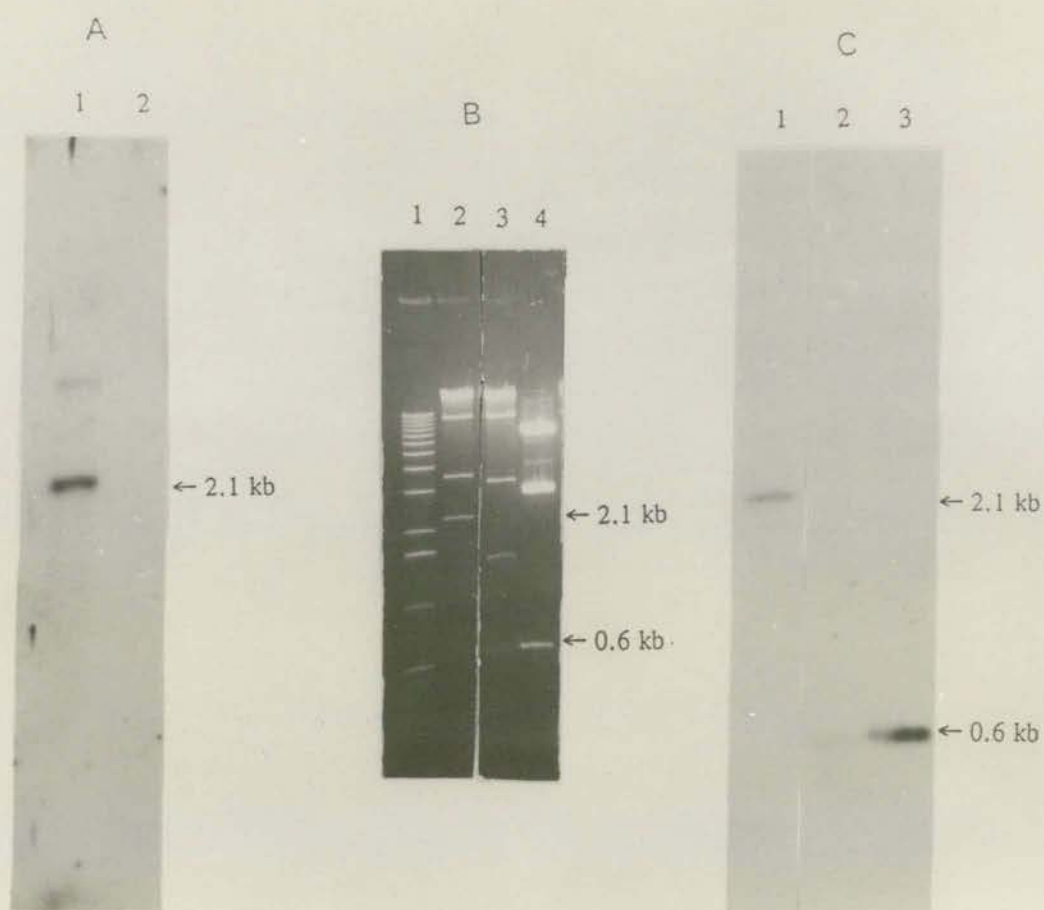
C. Southern blot of gel shown in B, probed with the 584 bp P-element

1 = EcoRI digestion

2 = EcoRI/HindIII digestion

3 = Carnegie 20, HindIII digestion

FIGURE 4.1



when plated on a hfl strain of E.coli such as NM514. The insertion of a foreign DNA fragment into the EcoRI or HindIII site disrupts the cI gene. Only the recombinant phage which are cI- can grow on the hfl host. Both the vector DNA of λ NM1149, and D19 genomic DNA were fully digested with EcoRI and ligated using T4 ligase at a molar ratio of 2:1 of vector:inserts. The ligation mixture was packaged in vitro using the λ -Packaging Kit (Amersham). The resulting library was titred on the bacterial strain NM514. The titre of the library was 1.6×10^6 pfu/ml.

50,000 recombinant phage were plated out on the host NM514 (23x23mm) and a lift was taken from the plate. The membrane lift was probed with the 584 bp HindIII P-element fragment labelled with digoxigenin-11-dUTP. This P-element fragment was also used to probe the Southern blot (Figure 4.1A), on which the P-element probe identifies the 2.1kb EcoRI fragment in the D19 flies.

Three positive clones were isolated after the secondary screen and named λ D19RI1, λ D19RI2 and λ D19RI3. DNA was digested with EcoRI, and with EcoRI and HindIII, and the digested fragments were electrophoresed. All the clones contain a 2.1 kb EcoRI fragment and a larger fragment (Figure 4.1B). A Southern blot of the gel was probed with a P-element fragment used to screen the genomic library. All three clones contain a 2.1 kb EcoRI fragment hybridizing to a P-element probe (data not shown). The results for λ D19RI1 are shown in figure 4.1C. In lane 1, the P-element probe detected a 2.1 kb EcoRI fragment corresponding to the 2.1 kb EcoRI fragment shown in the previous genomic Southern. In lane 2, the probe detected a EcoRI/HindIII fragment which has the size same as the 584 bp HindIII P-element fragment. These results show that the clone contains the 5'end of the transposon and a flanking *Drosophila* genome sequence. As in Fig 4.1B, there is an extra 5 kb EcoRI fragment in lane 1, which is probably caused by multiple ligation or by partial digestion of the genomic DNA. This fragment was not further analyzed.

4.3 Searching for Transcription Units

The assumption made in these experiment is that a nearby gene will show the same expression pattern as the *lac-Z* fusion gene. Thus a search for the transcription units flanking the transposon was carried out by Northern analysis.

The 2.1 kb EcoRI fragment from λ D19RI1 was used to probe a Northern blot containing ovarian RNA and female carcass RNA from Oregon R flies. Since no positive signals were detected (data not shown), further chromosome walking was pursued.

A genomic library of *D. melanogaster* constructed in λ Dash (a gift from Dr. R. Suanders) was used for the chromosome walking. Approximately 50,000 phage were plated on *E. coli* strain 646, which has the Spi selection for the recombinant. The phage plaques were transferred from the plate (23X23mm) onto a Hybond-N nylon membrane, and the membrane was probed with the radio-labelled 2.1 kb EcoRI fragment from λ D19RI. Seven positive clones were obtained. Figure 4.2A is the result of gel electrophoresis of the DNA from the six clones digested with EcoRI. All six clones contain a 11 kb EcoRI fragment and one or two smaller fragments. A Southern blot from the same gel was probed with the 2.3 kb EcoRI fragment from the λ D19RI. The 2.1 kb EcoRI fragment from λ D19RI hybridized to one of the small fragments of each clone as shown in Figure 4.2B. One of these λ Dash clones, referred to as λ D10, was chosen for further analysis because it has the biggest insert among the six. This clone contains EcoRI fragments of 11 kb, 5.5 kb and 0.2 kb.

In a preliminary experiment, the 5.5 kb and 11 kb EcoRI fragments from λ D10 were used to probe a Northern blot containing RNA extracted separately from ovaries, embryos, female carcasses and third instar larvae of Oregon R flies. While the 5.5 kb fragment failed to detect any RNA transcripts (data not shown), the 11 kb fragment hybridized at least four transcripts (Figure 4.3A). Among these transcripts is a 1.1 kb message, which is highly enriched in ovaries and embryos.

To obtain more information about each transcript, especially the 1.1 kb one detected by the λ D10 insert, the phage DNA was digested

Figure 4.2

A. Ethidium Bromide stained agarose gel containing EcoRI digested DNA

of λ Dash clones

- 1 = 1 kb ladder marker**
- 2 = λ D1**
- 3 = λ D3**
- 4 = λ D4**
- 5 = λ D7**
- 6 = λ D8**
- 7 = λ D9**
- 8 = λ D10**

C. Southern blot of gel shown in A, probed with the 2.1 kb EcoRI fragment from λ D19RI1

- 1 = λ D1**
- 2 = λ D3**
- 3 = λ D4**
- 4 = λ D7**
- 5 = λ D8**
- 6 = λ D9**
- 7 = λ D10**

The digestion was not complete, so signal can be seen in the phage arms.

FIGURE 4.2

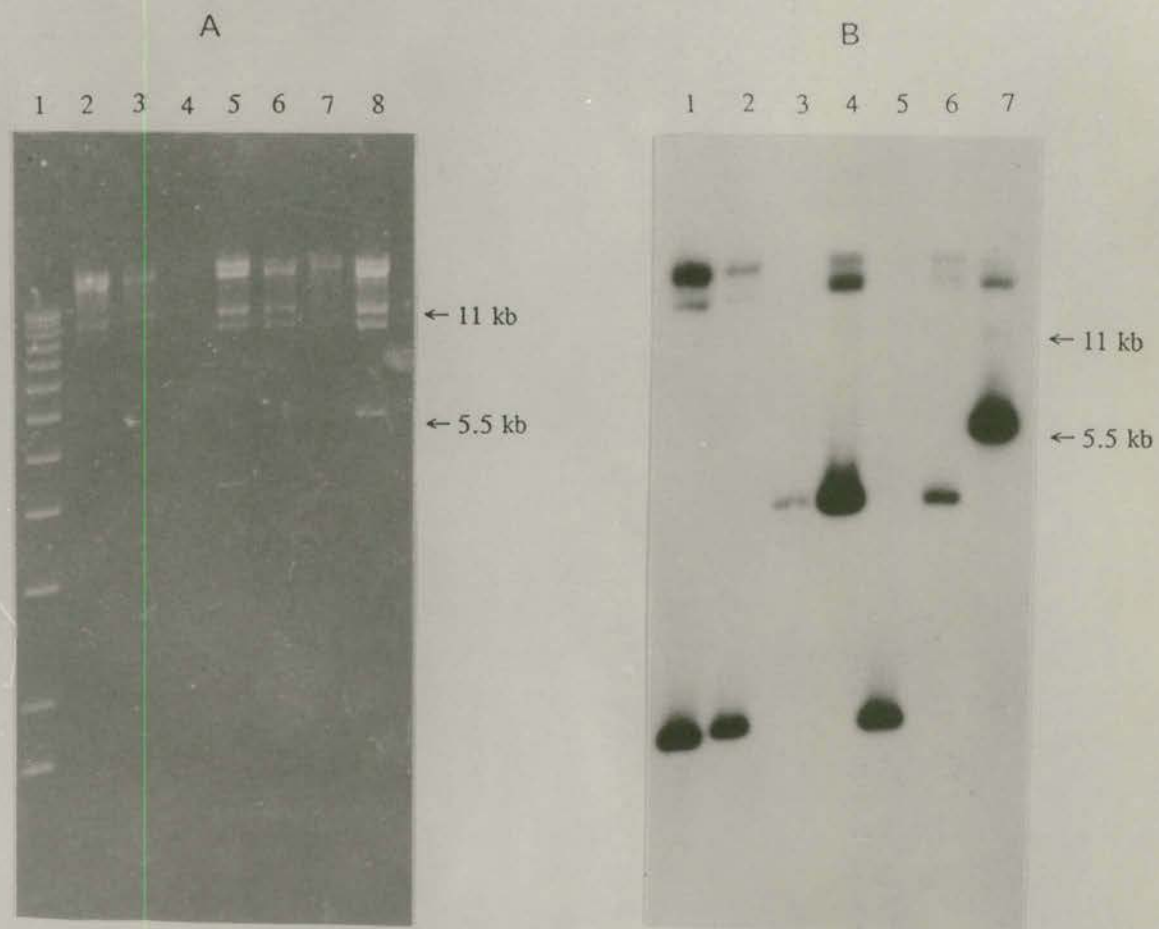


Figure 4.3

A. Northern blot probed with the 11 kb EcoRI fragment from λ D10
(see Figure 4.4).

- 1 = total RNA from female carcass
- 2 = total RNA from ovary
- 3 = total RNA from 3rd instar larvae
- 4 = total RNA from 0-6 hour embryos

This fragment detected multiple transcript, the 1.1 kb transcript was very abundant in ovary and embryo RNA.

B. Northern blots containing total male and female RNA, and reprobed with different HindIII fragments from λ D10.

- 1 = male RNA, GH1 probe
- 2 = female RNA, GH1 probe
- 3 = male RNA, GH2 probe
- 4 = female RNA, GH2 probe
- 5 = male RNA, GH3 probe
- 6 = female RNA, GH3 probe
- 7 = male RNA, GH4 probe
- 8 = female RNA, GH4 probe

C. Northern blot containing testis and ovary total RNA, probe with GH4 fragment from λ D10

- 1 = testis RNA
- 2 = ovary RNA

FIGURE 4.3

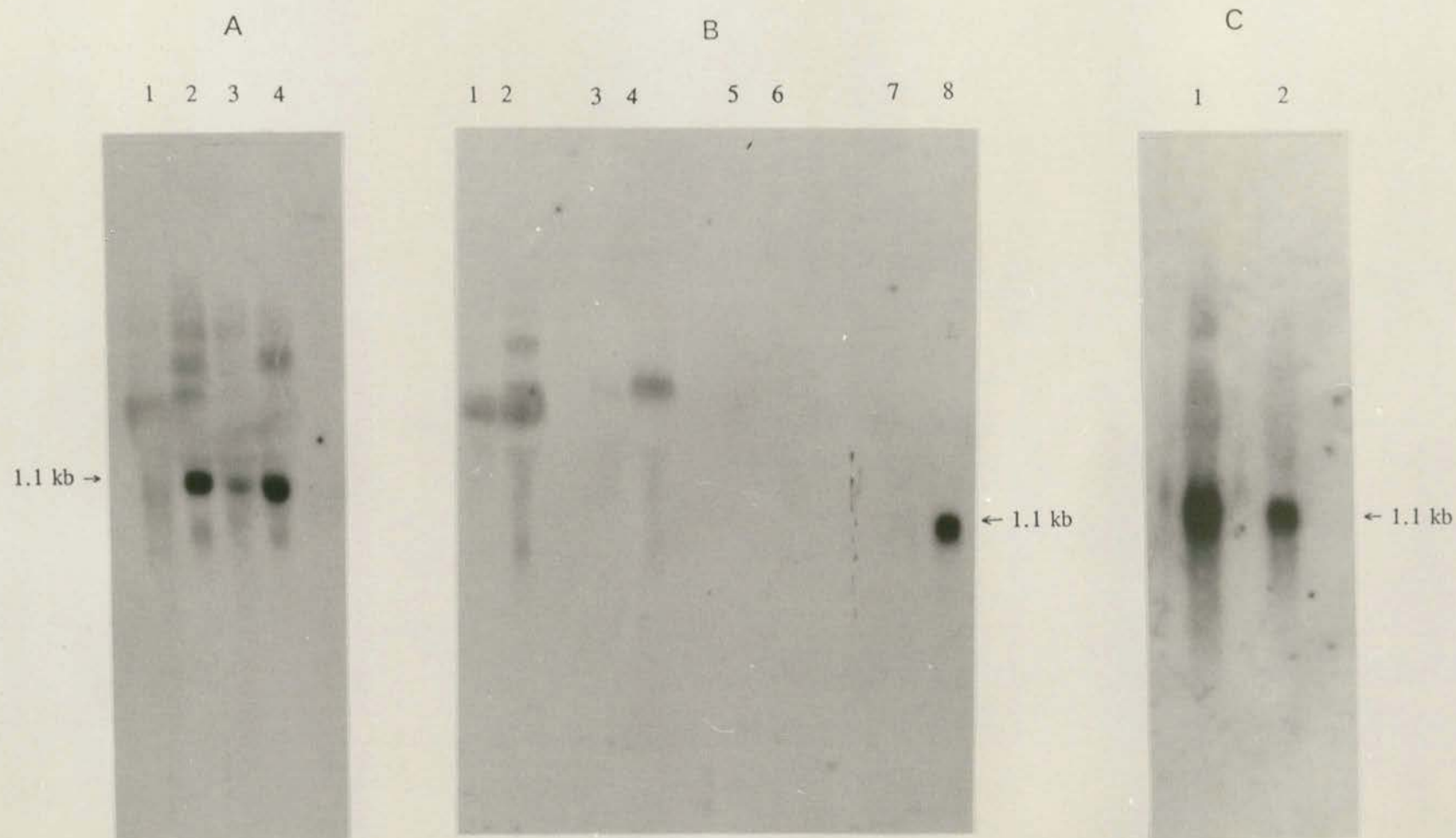
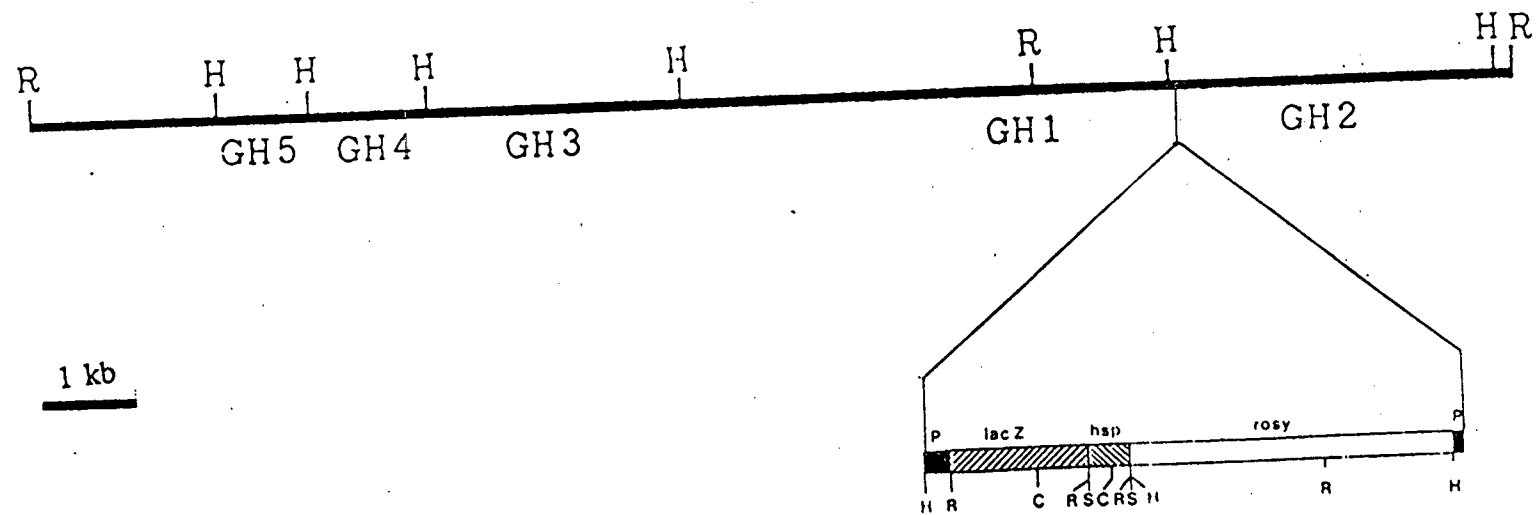


Figure 4.4

Restriction map of λ D10 showing EcoRI and HindIII sites. Phage arms are not shown. The position of the P[lac, ry⁺]A is illustrated.

Bar = 1kb

FIGURE 4.4



with HindIII. The insert of λ D10 contains five HindIII fragments. These are named as GH1 (5.5 kb), GH2 (3.2 kb), GH3 (2.5 kb), GH4 (1.3 kb) and GH5 (0.7 kb). Probes were made from four of these fragments (GH1-4) and used to probe Northern blots containing total RNA of male and female. The results are presented in Figure 4.3B. The GH1 probe detected two largest transcripts, the GH2 probe detected only one transcript, and the GH3 probe did not detect any positive signals. From the restriction map (Figure 4.4), GH2 is within the 5.5 kb EcoRI fragment, thus failure of the 5.5 kb fragment to detect any transcripts in previous Northern must due to poor labelling of that probe. The GH4 probe detected the abundant 1.1 kb transcript from female RNA. Since no transcripts were detected in the male RNA by GH4, a Northern blot containing ovary and testis RNA of Oregon R flies was probed with GH4 probe. An abundant 1.1 kb transcript was detected in both ovary and testis RNA (Figure 4.3C).

A restriction map was made for λ D10 with the restriction endonucleases EcoRI and HindIII (Figure 4.4). The positions of the P[lac, ry⁺] transposon is also shown in this map.

4.4 Discussion

By screening the genomic library constructed for the strain D19, a genomic sequence flanking the 5'end of the enhancer detector P[lac, ry⁺] has been cloned. A further 11 kb genomic fragment upstream of this genomic sequence and 4 kb fragment down stream of the transposon have been recovered by chromosome-walking. The 11 kb EcoRI fragment detected multiple transcripts. One female transcript was detected in female RNA by a 5 kb HindIII fragment (GH1) flanking the 5'end of the transposon. This indicates that the genomic sequence flanking the transposon might encode transcripts with a similar expression pattern to the *lacZ* fusion gene. A 1.1 kb message enriched in ovary, testis and embryo RNA was detected by a 1.3 kb HindIII fragment (GH4) 8 kb upstream of the transposon. In addition to its similar expression pattern to the *lacZ* fusion gene, this 1.1 kb transcript was particularly attractive for further study because of its high level expression in the ovary and

testis. These results support for the effectiveness of the "enhancer trap" technique for isolating genes expressed in the ovary.

The two genomic fragments GH1 and GH4 have been further analyzed and the results are presented in following chapters.

CHAPTER FIVE

ISOLATION AND SEQUENCING cDNAs CORRESPONDING TO THE GENOMIC FRAGMENT FLANKING THE P[LAC, RY⁺]A INSERTION

5.1 Introduction

As described in chapter 4, the 5.5 kb genomic fragment GH1 flanking the P[lac, ry⁺]A insertion which detected two ovary-specific transcripts, and the 1.3 kb genomic fragment GH4, which detected a 1.1 kb ovary and embryo enriched transcript, were used to screen a *Drosophila* ovary cDNA library. No cDNAs corresponding to the transcripts from the GH1 fragment were isolated, but two cDNAs corresponding to the transcript from the GH4 fragment were isolated. This chapter describes their isolation and characterization.

5.2 Screening of the Ovarian cDNA Library

An ovary cDNA library of *D. melanogaster* constructed in the vector gt11 (a gift from P. Sullivan and L. Kalfayan, University of North Carolina) was plated on a BBL-amp plate with the host cell Y1090. 50,000 recombinant phage were screened using the 1.3 kb fragment GH4 of λ D10 as probe (this detected the 1.1 kb ovary enriched transcript). After secondary screening, two positive clones were obtained. These are called λ ovD19C1 and λ ovD19C2. Each clone contains a single EcoRI insert; λ ovD19C1 contains an EcoRI fragment of 1.8 kb and λ ovD19C2 contains an EcoRI fragment of 1 kb. The two cDNAs were then subcloned into the EcoRI site of pBluescript SK⁺ and referred to as pBD19C1 for the 1.8 kb cDNA and pBD19C2 for the 1 kb cDNA. The 1.8 kb cDNA was rather larger than expected for a cDNA encoding a 1.1 kb transcript.

5.3 Analysis of the Homology of the cloned cDNA to the Genomic clone

5.3.1 Southern Analysis

To confirm the homology of the two cDNAs to λ D10, probes made separately from ov19C1 and ov19C2 were used to hybridise to Southern blots containing EcoRI and Hind III-digested λ D10 DNA and genomic DNA of Oregon-R flies.

When the 1.8 kb ovD19C1 fragment was used to probe Southern blots containing HindIII- and EcoRI-digested λ D10, it only hybridized to a 1.3 kb HindIII fragment representing GH4 (Figure 5.1). As seen from the restriction map of λ D10 (Figure 4.4), GH4 lies in the middle of the λ D10 insert, and thus ovD19C1 must contain

Figure 5.1

A. Ethidium Bromide stained agarose gel showing the EcoRI and HindIII

digested DNA from λ D10.

1 = 1 kb ladder maker

2 = EcoRI digested λ D10 DNA

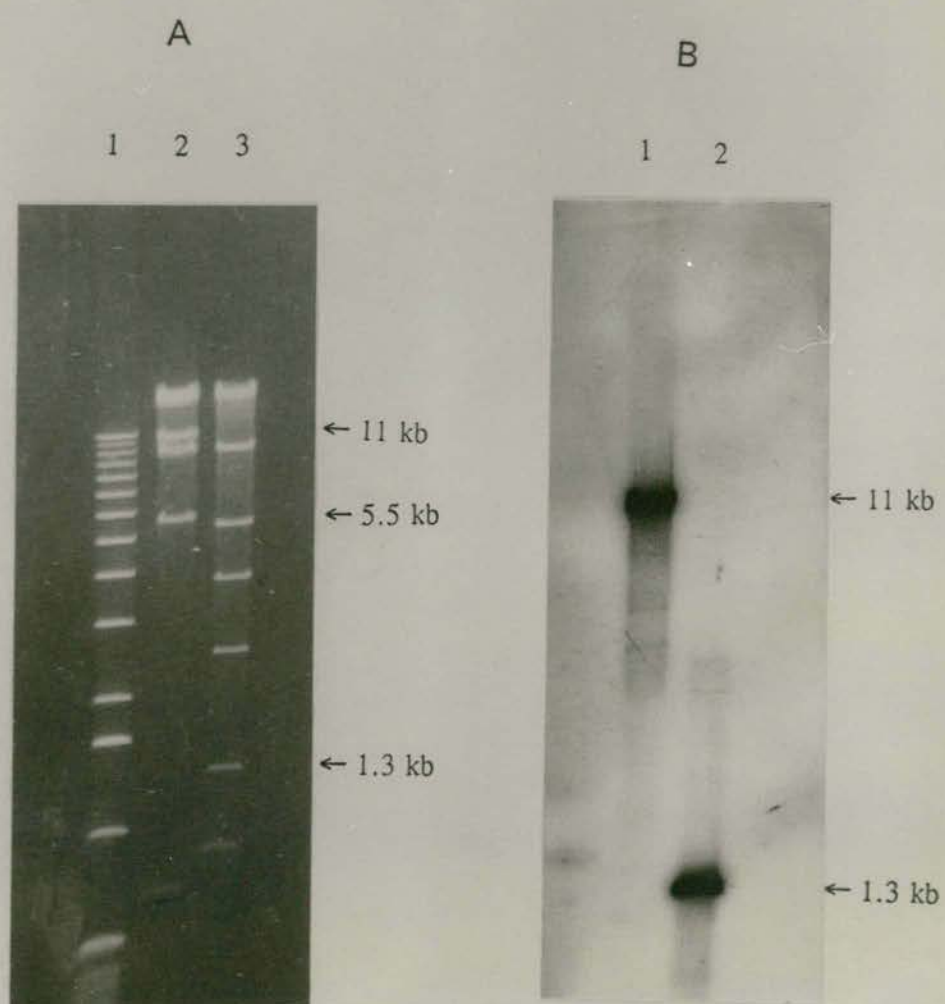
3 = HindIII digested λ D10 DNA

B. Southern blot from the gel shown in A, and probed with ovD19C1 and ovD19C2. The filter was first probed with ovD19C1, then the probe was stripped and re-probed with ovD19C2. This picture shows the result of ovD19C2 probe.

1 = EcoRI digestion

2 = HindIII digestion

FIGURE 5.1



DNA sequences which are different from λ D10. On a Southern of genomic DNA of Oregon R, as well as detecting a 1.3 k HindIII fragment, ovD19C1 hybridized to a 11 kb HindIII fragment (Figure 5.2A). From the restriction map of λ D10, the 11 kb HindIII fragment detected on the genomic Southern must lie 12 kb or 3 kb outside the λ D10 insert. It was possible that there may be a large intron in the genomic sequence between two exons of ovD19C1. This was ruled out by subsequent experiments, as both in situ hybridization to the polytene chromosomes (5.3.1) and DNA sequencing analysis show that this cDNA is actually a cloning artefact representing two mRNAs. Since there are no EcoRI sites within ovD19C1 other than the two at the ends of this cDNA, it is likely that the ligation of the two molecules happened during the process of adding the EcoRI linkers.

The ovD19C1 probe was stripped off from the Southern blot containing the λ 10 DNA described above and re-probed with ovD19C2. The ovD19C2 probe detected a single 1.3 kb HindIII fragment and a single 11 kb EcoRI fragment (Figure 5.1). On a genomic Southern, the ovD19C2 probe only detected an 11 kb EcoRI fragment and the 1.3 kb HindIII fragment as expected (Figure 5.2B).

5.3.2 *In Situ* Hybridization to the Third Instar larval Polytene Chromosomes

The chromosomal locations of DNA fragments of interest have been very informative for both genetic and molecular studies. For this project, it was important to obtain the chromosomal locations of both the transposon and the two cloned cDNAs for the following reasons. 1. This would confirm whether the cDNAs map to the same position as the transposon, as predicted. 2. As described in 4.1, the 1.8 kb cDNA was longer than predicted. If it was a cloning artefact caused by the ligation of two molecules, they would map to different locations on the polytene chromosomes. 3. It provides important information for future mutagenesis experiments either by imprecise excision of the transposon, or by conventional EMS mutagenesis or P-element mediated mutation.

Figure 5.2

A. Southern blot containing genomic DNA from Oregon R flies, digested with EcoRI, HindIII and SalI, probed with ovD19C1. The DNA was not completely cut.

1 = EcoRI digestion

2 = HindIII digestion

3 = SalI digestion

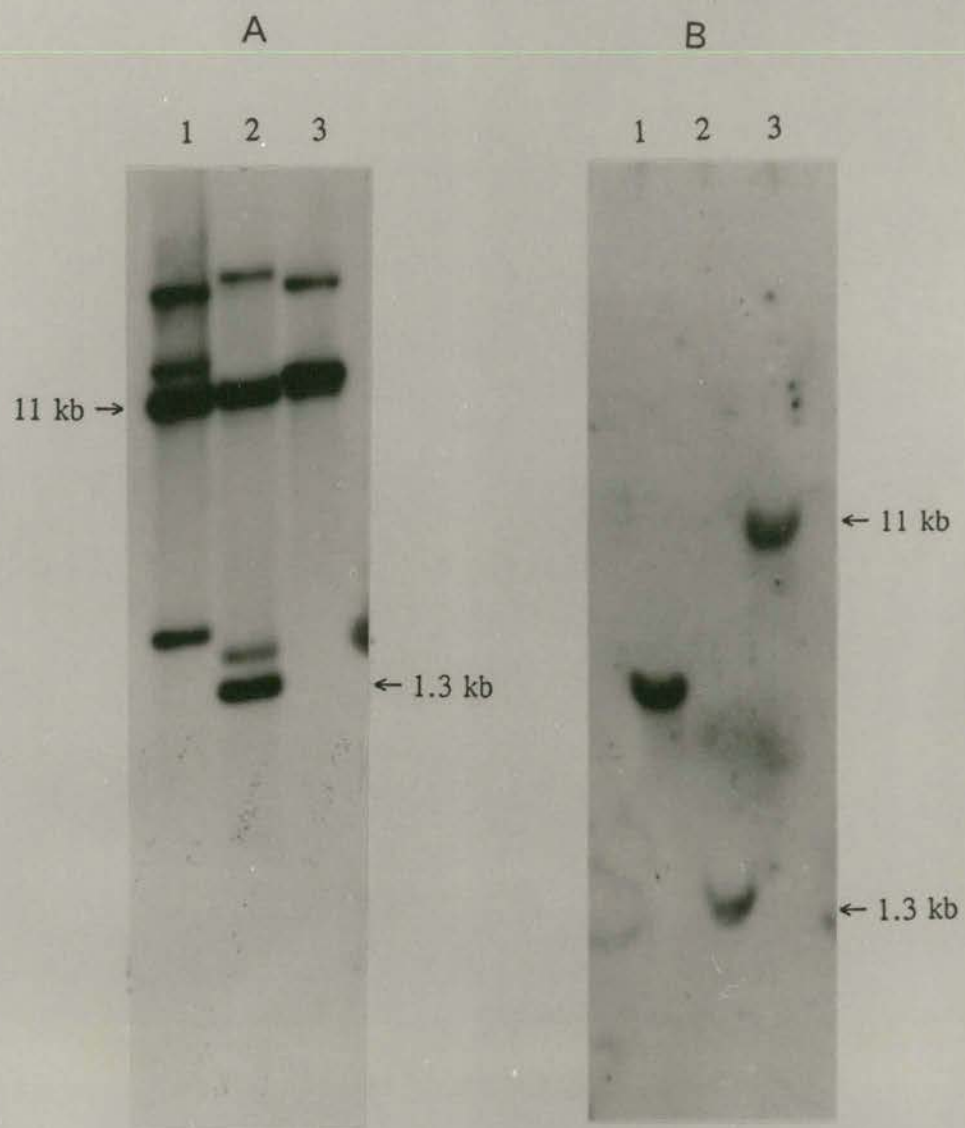
B. Southern blot containing genomic DNA from Oregon R flies, digested with BamHI, EcoRI and HindIII, probed with ovD19C2.

1 = BamHI digestion

2 = HindIII digestion

3 = EcoRI digestion

FIGURE 5.2



Methods for chromosome preparation, hybridization and signal detection are described in 2.5.14. Third instar larvae of the D19 strain and Oregon R strain were used.

Mapping the P-transposon P[lac, ry[±]]

To find the chromosomal location of the transposon P[lac, ry⁺], a digoxigenin-labelled *lacZ* RNA probe was made by in vitro transcription using the T7 RNA polymerase from pGem-1-lacZB. This plasmid contains the lac-Z gene under the control of bacterial phage T7 and SP6 promoters in opposite directions.

The P[lac, ry⁺] insert maps to 98F on the right arm of chromosome 3 (Figure 5.3A)(confirmed by Dr. K. Lineruth)

Mapping ovD19C1 and ovD19C2

For mapping ovD19C1 and ovD19C2, RNA probes were made in the same way as the *lacZ* probe, but using pBD19C1 and pBD19C2 as templates. For both plasmids T7 promoter was used for transcription.

The probes made from pBD19C1 and pBD19C2 both detected a signal at the same location as that found for the transposon at 98F (Figure 5.3B). But ovD19C1 hybridized to a band on the left arm of chromosome 3 (Figure 5.3C). This helped to explain why ovD19C1 detected two EcoRI and HindIII fragments on the genomic Southern while only detecting one EcoRI and HindIII fragment in λ D10 DNA as described above. It appears that ovD19C1 is, in fact, two molecules which have been ligated during the construction of the cDNA library.

5.4 Sequencing Analysis of GH4 and the Two cDNAs

5.4.1 Sequencing Strategy

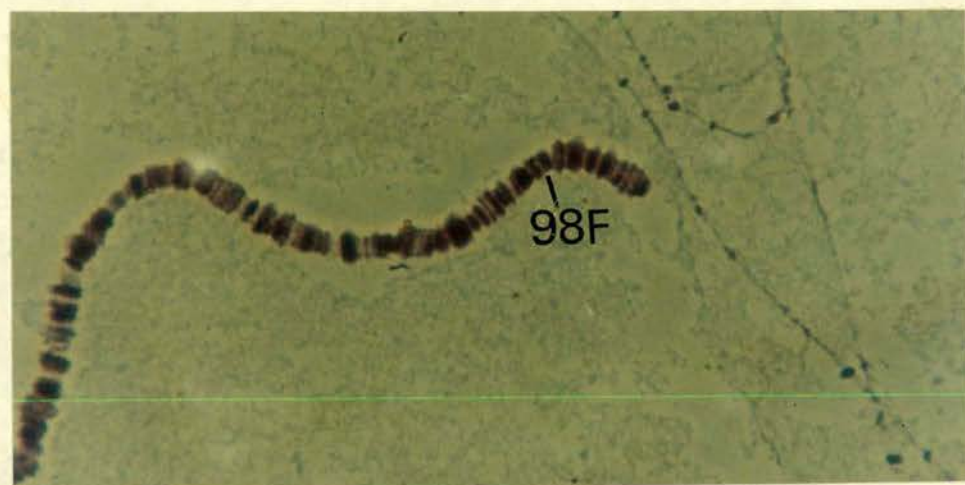
Nested deletions were generated for pBD19GH4 and the two cDNA clones pBD19C1 and pBD19C2 using Exonuclease III and nuclease S1 which produces a series of unidirectional deletions for each plasmid suitable for sequencing. Deletions were made from both directions of pBD19GH1 and pBD19C2. Only one direction of deletions were made for pBD19C1 starting from the KpnI site.

Figure 5.3

In situ hybridization to 3rd instar larval polytene chromosomes

- A. Chromosome preparation from D19 strain, probed with *lacZ* probe
- B. Chromosome preparation from Oregon R strain, probed with ovD19C2 probe
- C. Chromosome preparation from Oregon R strain, probed with ovD19C1 probe

FIGURE 5.3



A



B



C

Synthesized oligo-nucleotide primers were used to connect some deletions where the gap between two deletions is beyond 300bp. DNA sequencing was performed by the dideoxynucleotide-chain-termination method (2.5.13) using the T7 DNA polymerase. Both the "T7 Sequencing Kit" (Pharmacia) and home made reagents were used. The vector pBluescript II SK⁺ contains the f1 origin of replication from the f1 filamentous phage, allowing rescue of single stranded DNA upon coinfection with helper phage. Single stranded DNA templates were rescued from the "plus" strand of pBluescript II SK⁺ for each plasmid with the helper phage VCM13 while the opposite strand of each template was sequenced using the NaOH denaturing method (2.5.13). Sequencing strategies are shown in Figure 5.4.

5.5 Sequence Analysis of OvD19C2

5.5.1 Sequencing of OvD19C2

Both strands of ovD19C2 DNA in the plasmid pBD19C2 were sequenced at least twice. The whole sequence of this cDNA is presented in Figure 5.5. The ovD19C2 contains 1038 nucleotides. In addition, there is a 44 bp sequence in ovD19C1 which extends from the 5'end of ovD19C2 and which has an identical sequence to GH4. This 44 bp sequence is added to ovD19C2, thus making it 1082 bp in length. Compared to the size of the 1.1 kb transcript detected by GH4, it appears that this cDNA is almost full length. The presence of a 60bp poly(A) tail which lies 34bp down stream of the polyadenylation signal ATTAAA defines the 3'end of this cDNA and helps to determine the sense and anti-sense strand.

The cDNA sequence was translated into a polypeptide sequence using the "MAP" program in the GCG7 package. The longest open-reading frame (ORF) from this cDNA initiates at the first ATG codon in the sequence at nucleotide 102 and ends at an in frame stop codon TAG at position 783. This ORF contains 681 bp nucleotides encoding a predicted protein with 227 amino acid residues (Figure 5.5). The estimated molecular weight of this protein is 25 kD. The AAAAATG initiation sequence at the beginning of the ORF matches to the C/A-AA-A/C-AUG consensus of translation initiation sites

Figure 5.4
Sequencing strategies

A: ovD19C2

B: GH4 and its 3' and 5' flanking region

C ovD19C1

FIGURE 5.4

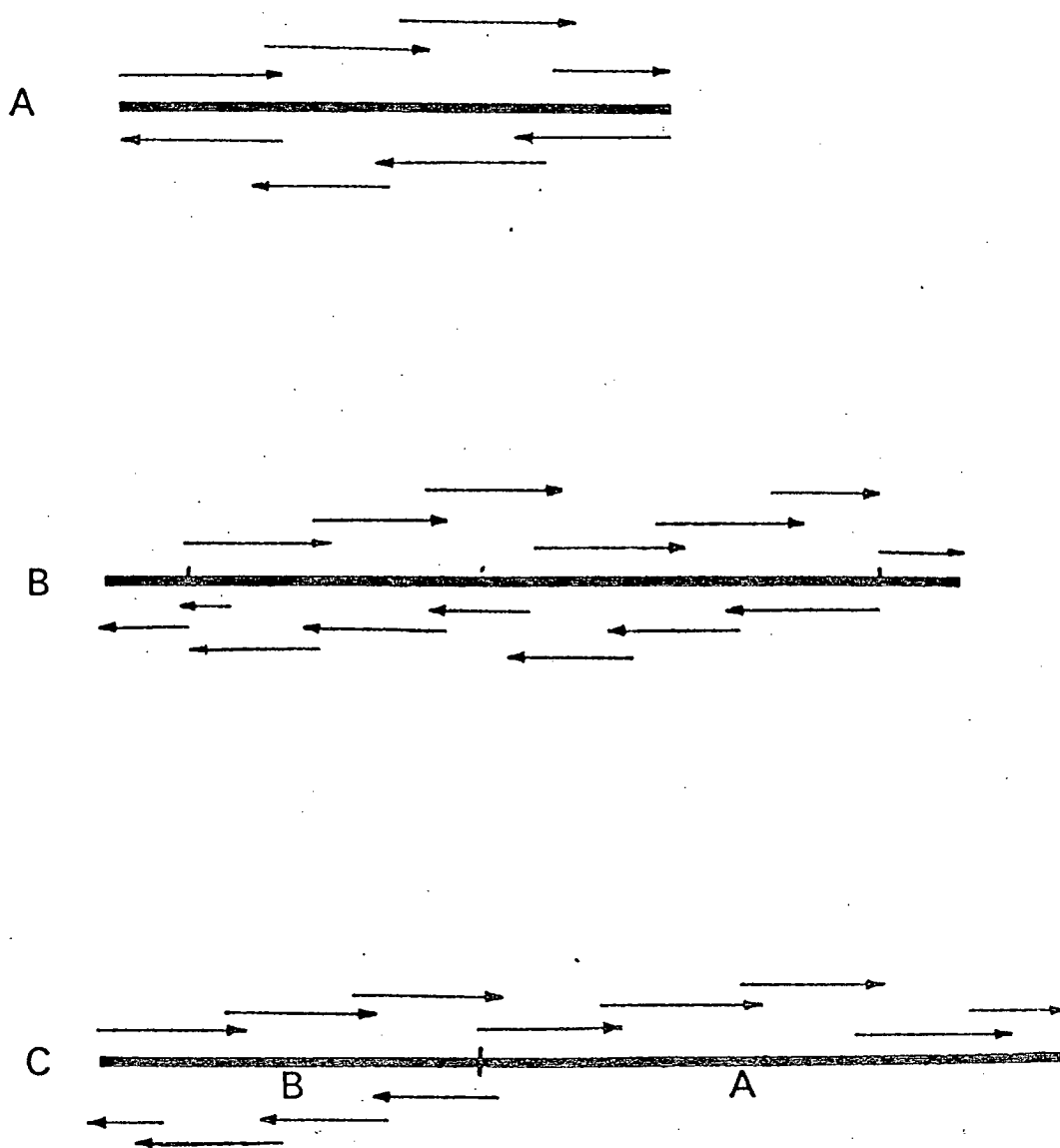


Figure 5.5

DNA sequence of ovD19C2, and the amino acid sequence predicted from the longest ORF of this cDNA. The first 44 nucleotides at the 5'end is derived from ovD19C1 (5.4.1.1). The ATG translation initiation codon, and the polyadenylation signal sequence are underlined

Figure 5.5 DNA sequence of ovd19C2 and its predicted amino acid sequence

CCACATTCAGAAAAAAGTAAAGCTCGTCGTAATTGTGTGTAGAAATCCTTCATTTTC
CTGTGAATTTTCAATCAGATCACTGCTAAAGCCAAATTCAAAAATATGTTAACCTGG
MetLeuThrTrp
ACGCCACTTGAATCTAATCCCGAGGTTTTGACCAAGTACATACATAAACTGGGCGTG
ThrProLeuGluSerAsnProGluValLeuThrLysTyrIleHisLysLeuGlyVal
TCGCCAGCCTGGTCCGTAACTGACGTCATTGGATTGGAAGATGACACCTTGGAAATGG
SerProAlaTrpSerValThrAspValIleGlyLeuGluAspAspThrLeuGluTrp
ATTCCGCGTCCCGTAAAGGCGTTTCAATTTTGCTCTTCCCGTGCAGCGAAACCTATGAG
IleProArgProValLysAlaPheIleLeuLeuPheProCysSerGluThrTyrGlu
AAGCACCGCACCGAGGAGCACGATCGGATTAAGGAGGTGGAGGAGCAGCATCCCGAG
LysHisArgThrGluGluHisAspArgIleLysGluValGluGluGlnHisProGlu
GATCTCTTCTACATGCGCCAGTTCACCCACAACGCCTGTGGAACCGTCGCCCTGATC
AspLeuPheTyrMetArgGlnPheThrHisAsnAlaCysGlyThrValAlaLeuIle
CACAGCGTGGCCAACAACAAAGAAGTGGACATTGACCGCGGAGTACTGAAGGACTTC
HisSerValAlaAsnAsnLysGluValAspIleAspArgGlyValLeuLysAspPhe
CTGGAGAAGACAGCTTCCTTGTCCCCGAGGAACGCGGACGGGCCCTCGAAAAAGAC
LeuGluLysThrAlaSerLeuSerProGluGluArgGlyArgAlaLeuGluLysAsp
GAGAAGTTCACCGCCGATCATGAGGCCTTGGCTCAAGAGGGCCAGACGAATGCCGCC
GluLysPheThrAlaAspHisGluAlaLeuAlaGlnGluGlyGlnThrAsnAlaAla
AATCATGAGAAGGTGATCCACCACTTCATCGCACTGGTGAACAAGGAGGGTACTCTG
AsnHisGluLysValIleHisHisPheIleAlaLeuValAsnLysGluGlyThrLeu
TACGAGCTGGATGGCCGCAAGTCCTTCCCGATCAAGCACGGACCGACTTCGGAGGAG
TyrGluLeuAspGlyArgLysSerPheProIleLysHisGlyProThrSerGluGlu
ACTTTTGTGAAGGATGCGGCCAAGGTCTGCAAGGAGTTCATGGCTCGCGATCCCAAC
ThrPheValLysAspAlaAlaLysValCysLysGluPheMetAlaArgAspProAsn
GAAGTGCCTTCACCGTTTTGGCCTTGACCGCGGCACAACAATAGGGTGGAACTAAT
GluValArgPheThrValLeuAlaLeuThrAlaAlaGlnGln
CCCATAACTGAGAACCACCAAGATCCCGGATCTGCCAGCTATATATTTATATATCCA
ACATATTTTAAACCCAGGCTGCAAAGGCGAACGAAGCCCGCCTTTCAGCCGGTCAAT
AGTTTACGCATATTTATTCCGCATTGACCCAAAGCCCCATATGTATTGGGAACACCA
TGATTGTTATTAAGAGTGCACTCTAGCTTGATGCTAAGCTTATACTGAAAAAAAAA
AAA

described for *Drosophila* (Cavener, 1987). The codon usage in this ORF is typical to other *Drosophila* genes (see below). The longest ORF in the two other frames are only 216 bp and 189 bp.

5.5.2 Codon Usage of OvD19C2

There is a bias in the choice of degenerated codons in many genes. Synonymous codon choice is usually similar among the taxonomically related organisms. In mammals C and G are often used as the third codon base, whereas A and T are preferred in viruses (Grantham *et al.* 1980). Inspection of 17 *Drosophila* genes shows that the preference of the third codon base in these genes is similar to that of mammals; C and G are most frequently used as the third codon base in *Drosophila*. To determine whether or not the codon usage in the longest ORF of ovD19C2 has the same bias as other *Drosophila* genes, a codon frequency survey was conducted. The result is presented in Table 5.1. This table shows the codon usage in the predicted protein of ovD19C2 is consistent with previously published data.

5.5.3 Comparison of ovD19 cDNA sequences Genomic sequence

The two strands of GH4 were sequenced, and shows that GH4 contains 1300bp. 150 bp sequence of GH3 which flanks the 5'end of GH4, and 150 bp sequence of GH2 flanking 3'end of GH4 were determined. The junction between GH3 and GH4 was sequenced by using λ D10 as the template and using a synthetic oligonucleotide primer priming from 150 bp at the 5'end of GH4. All together, 1631 bp of genomic sequence has been determined. It has been called GenD19.

A comparison between the genomic sequence and ovD19C2 cDNA sequence was made and is presented in Figure 5.6. This shows that GH4 completely covers the 5'end of ovD19C2 while at the 3'end GH4 ends at a HindIII site is only 5 bp upstream of the poly(A) tail of ovD19C2. Although the transcription initiation site of ovD19C2 has not been experimentally determined, a TATAA sequence is found at 33bp upstream of ovD19C2. Also a CAAT box is found 106 bp upstream of this cDNA.

Table 5.1

Codon Frequency of ovD19C2

| AmAcid | Codon | Number | /1000 | Fraction |
|--------|-------|--------|-------|----------|
| Gly | GGG | 0.00 | 0.00 | 0.00 |
| Gly | GGA | 5.00 | 22.03 | 0.56 |
| Gly | GGT | 1.00 | 4.41 | 0.11 |
| Gly | GGC | 3.00 | 13.22 | 0.33 |
| Glu | GAG | 19.00 | 83.70 | 0.70 |
| Glu | GAA | 8.00 | 35.24 | 0.30 |
| Asp | GAT | 7.00 | 30.84 | 0.54 |
| Asp | GAC | 6.00 | 26.43 | 0.46 |
| Val | GTG | 8.00 | 35.24 | 0.50 |
| Val | GTA | 3.00 | 13.22 | 0.19 |
| Val | GTT | 2.00 | 8.81 | 0.13 |
| Val | GTC | 3.00 | 13.22 | 0.19 |
| Ala | GCG | 3.00 | 13.22 | 0.16 |
| Ala | GCA | 2.00 | 8.81 | 0.11 |
| Ala | GCT | 3.00 | 13.22 | 0.16 |
| Ala | GCC | 11.00 | 48.46 | 0.58 |
| Arg | AGG | 0.00 | 0.00 | 0.00 |
| Arg | AGA | 0.00 | 0.00 | 0.00 |
| Ser | AGT | 0.00 | 0.00 | 0.00 |
| Ser | AGC | 2.00 | 8.81 | 0.22 |
| Lys | AAG | 14.00 | 61.67 | 0.82 |
| Lys | AAA | 3.00 | 13.22 | 0.18 |
| Asn | AAT | 3.00 | 13.22 | 0.38 |
| Asn | AAC | 5.00 | 22.03 | 0.63 |
| Met | ATG | 3.00 | 13.22 | 1.00 |
| Ile | ATA | 1.00 | 4.41 | 0.10 |
| Ile | ATT | 5.00 | 22.03 | 0.50 |
| Ile | ATC | 4.00 | 17.62 | 0.40 |
| Thr | ACG | 2.00 | 8.81 | 0.12 |
| Thr | ACA | 1.00 | 4.41 | 0.06 |
| Thr | ACT | 4.00 | 17.62 | 0.24 |
| Thr | ACC | 10.00 | 44.05 | 0.59 |
| Trp | TGG | 3.00 | 13.22 | 1.00 |
| End | TGA | 0.00 | 0.00 | 0.00 |
| Cys | TGT | 1.00 | 4.41 | 0.33 |
| Cys | TGC | 2.00 | 8.81 | 0.67 |
| End | TAG | 0.00 | 0.00 | 0.00 |
| End | TAA | 0.00 | 0.00 | 0.00 |
| Tyr | TAT | 1.00 | 4.41 | 0.25 |
| Tyr | TAC | 3.00 | 13.22 | 0.75 |

Table 5.1 continued

| AmAcid | Codon | Number | /1000 | Fraction |
|--------|-------|--------|-------|----------|
| Leu | TTG | 8.00 | 35.24 | 0.40 |
| Leu | TTA | 1.00 | 4.41 | 0.05 |
| Phe | TTT | 1.00 | 4.41 | 0.09 |
| Phe | TTC | 10.00 | 44.05 | 0.91 |
| Ser | TCG | 3.00 | 13.22 | 0.33 |
| Ser | TCA | 0.00 | 0.00 | 0.00 |
| Ser | TCT | 1.00 | 4.41 | 0.11 |
| Ser | TCC | 3.00 | 13.22 | 0.33 |
| Arg | CGG | 2.00 | 8.81 | 0.20 |
| Arg | CGA | 0.00 | 0.00 | 0.00 |
| Arg | CGT | 1.00 | 4.41 | 0.10 |
| Arg | CGC | 7.00 | 30.84 | 0.70 |
| Gln | CAG | 3.00 | 13.22 | 0.50 |
| Gln | CAA | 3.00 | 13.22 | 0.50 |
| His | CAT | 4.00 | 17.62 | 0.36 |
| His | CAC | 7.00 | 30.84 | 0.64 |
| Leu | CTG | 7.00 | 30.84 | 0.35 |
| Leu | CTA | 0.00 | 0.00 | 0.00 |
| Leu | CTT | 1.00 | 4.41 | 0.05 |
| Leu | CTC | 3.00 | 13.22 | 0.15 |
| Pro | CCG | 5.00 | 22.03 | 0.45 |
| Pro | CCA | 2.00 | 8.81 | 0.18 |
| Pro | CCT | 0.00 | 0.00 | 0.00 |
| Pro | CCC | 4.00 | 17.62 | 0.36 |

Figure 5.6

DNA sequence of the genomic region covering ovD19C2. The sequences homologous to ovD19C2 are in upper case, others are in lower case. Introns are underlined.

The intron splicing sequences are shown at the end of this figure.

Figure 5.6 Genomic Sequences of GenD19

Letters in upper case show homologous to ovD19C2, those in lower case show 5' and 3' region flanking ovD19C2. Introns are in lower case and underlined. CAT box and TATA box are also underlined.

```

1  taatttttatt tccttttcct atttcgcatt cttgtacgaa tcaattttct
    -109
51  tctaataaat caccctaccc aaccaattaa gattaaatgc ttaacatttc
    GH4⇒ -38
101 tatttaaaaa ctataacaaa ctgataaatt gatgcgatta agcttataat
    +1
151 cgatatgttt cgaccaaadc aatgttcaca caCCACATTC AGAAAAAGT
201 AAAGCTCGTC AAAAGTGCCG GTAATTGTGT GTAGAAATCC TTCATTTTCC
251 TGTGAATTTT CAATCAGATC ACTGCTAAAG CCAAAATTC AAATATGTTA
    +149
301 ACCTGGACGC CACTTGAATC TAATCCCGAG gtaatgactt tacctgtaaa
351 ggtttttgcct accgcacttc cagttttttgc acattcattg tttctttggt
401 ccaaaatgtc ggattttggat aataaagtta aaagacgcca ttaagacctt
    +285
451 ccattccact ttcacagGTT TTGACCAAGT ACATACATAA ACTGGCGGTG
501 TCGCCAGCCT GGTCCGTAAC TGACGTCATT GGATTGGAAG ATGACACCTT
551 GGAATGGATT CCGCGTCCCG TAAAGGCGTT CATTTTGCTC TTCCCGTGCA
    +426
601 GCGAAACCGt aggttttagct tgaaccttct tacttattaa tcctggccaa
    +487
651 ccaggattcc aatttttagT ATGAGAAGCA CCGCACCGAG GAGCACGATC
701 GGATTAAGGA GGTGGAGGAG CAGCATCCCG AGGATCTCTT CTACATGCGC
751 CAGTTCACCC ACAACGCCTG TGAACCGTC GCCCTGATCC ACAGCGTGCC
    +631
801 CAACAACAAA GAGTGAgttc gtgacgtttt ttctatataa gatcgttaat
    +710
851 aatcgatatct ctttatatgt cacaaaacta atctttcccc agAGTGAGACA
901 TTGACCGCGG AGTACTGAAG GACTTCCTGG AGAAGACAGC TTCCTTGTCC
951 CCGGAGGAAC GCGGACGGGC CCTCGAAAAA GACGAGAAGT TCACCGCCGA
1001 TCATGAGGCC TTGGCTCAAG AGGGCCAGAC GAATGCCGCC AATCATGAGA
1051 AGGTGATCCA CCACTTCATC GCACTGGTGA ACAAGGAGGG TACTCTGTAC
1101 GAGCTGGATG GCCGCAAGTC CTTCCCGATC AAGCACGGAC CGACTTCGGA
1151 GGAGACTTTT GTGAAGGATG CGGCCAAGGT CTGCAAGGAG TTCATGGCTC

```

1201 GCGATCCCAA CGAAGTGCGC TTCACCGTTT TGGCCTTGAC CGCGGCACAA
1251 CAATAGGGTG GAACTAATCC CATAACTGAG AACCACCAAG ATCCCGGATC
1301 TGCCAGCTAT ATATTTATAT ATCCAACATA TTTTAAACCC AGGCTGCAAA
1351 GGCGAACGAA GCCCGCCTTT CAGCCGGTCA ATAGTTTACG CATATTTATT
1401 CCGCATTGAC CCAAAGCCCC ATATGTATTG GGAACACCAT GATTGTTATT
 end of GH4
1451 AAAGAGTGCA CTCTAGCTTG ATGCTAAAGCT TATACTgaac accacacggc
1501 ttttacttat taatatacat atataaaaat ataggcttggt ctggttcaaac
1551 tccccctttg tccaccacct ctgtgccaat agctgataac gatgtggcct
1601 ccattgcgca tgccctgccca attagtgtgc t

Intron 1: ATCCCGAGgtaatg.....ctttcacagGTTTTGAC
Intron 2: GCGAAACCgtaggt.....aatttttagTATGAGAA
Intron 3: AAGAGTGAgttcgt.....tttcccagAGTGGACA

CGgtgagt.....tttttttntagG
A a cccccccc c

The comparison between the genomic sequence and the sequence of ovD19C2 has revealed three discontinuous sequences in the genomic DNA. These discontinuous sequences represent the introns in this gene. The three introns are located at positions +149 to +285, +426 to +487 and +613 to +710, respectively (Figure 5.6). All three introns contain the *Drosophila* consensus sequences for splicing nuclear mRNA introns (Keller and Noon, 1985). The three intron splicing sequences and the *Drosophila* consensus sequences for splicing are listed in Figure 5.6.

5.5.4 Computer Data Base Searching

A. Protein Level

The protein sequence from the longest ORF of ovD19C2 was used to search the protein data base Prosrch Version 1.1 (Collins and Coulson, personal communication). It shows a high similarity to the human neuron-specific protein PGP 9.5. (Day and Thompson, 1987); the two proteins have 99 identical amino acid residues and 87 conserved changes. When the amino acid sequence of ovD19C2 was used to search the protein data base Swissprot, it shows similarity to the human ubiquitin carboxyl hydrolase isozyme L3 (UCH-L3) and isozyme L1 (Wilkinson *et al.*, 1989). The ovD19C2 protein shares 51.1% identity in 225 amino acid overlaps to human UCH-L3, with 115 identical residues and 77 conserved changes. Wilkinson and his colleagues showed that the human neuron specific protein PGP9.5 is the ubiquitin carboxyl terminal hydrolase isozyme L1 (Wilkinson *et al.*, 1989). Figure 5.7 shows the comparison between the putative protein of ovD19C2 and the two human proteins. This uses the FastA program which gives the percentage of identity and the identical residues as well as the conserved substitutions between two protein sequences. Miller *et al.* (1989) have cloned the yeast ubiquitin hydrolase gene, and the protein expressed from the yeast cDNA shows the ubiquitin hydrolase activity both in vivo and in vitro. Figure 5.8 is a alignment of the two human proteins UCH-L3 and PGP 9.5 and the yeast protein with the putative ovD19C2 protein. The features defining this class of enzymes will be dealt in 5.4.3.

Figure 5.7

A comparison of the amino acid sequences between ovD19C2 protein and the human ubiquitin carboxyl terminal hydrolase isozyme L3 (h-UCH-L3) and isozyme L1 (PGP 9.5) using the program FastA.

Identical amino acid residues between two sequences are indicated by lines, the conserved substitutions are indicated by two dots.

A. ovD19C2 to h-UCH-L3

1 = amino acid sequence of ovD19C2

2 = amino acid sequence of h-UCH-L3

B. ovD19C2 to PGP 9.5 (h-UCH-L1)

1 = amino acid sequence of ovD19C2

2 = amino acid sequence of PGP 9.5 (h-UCH-L1)

Figure 5.7A

SCORES Init1: 282 Initn: 583 Opt: 621
51.1% identity in 225 aa overlap

```

      10      20      30      40      50
1  MLTWTPLESNPEVLTKYIHKLGVS PAWSVTDVIGLEDDTLEWIPRPVKAFILLFPCSE
   | |||:|||| :|||:|:| :|| :|| :| :||| :||| :|
2  MEGQRWLPLEANPEVTNQFLKQLGLHPNWQFVDVYGMDPELLSMVPRPVCAVLLLPITE
      10      20      30      40      50      60

      60      70      80      90      100     110
1  TYEKHRTTEHDRIKEVEEQHPEDLFYMRQFTHNACGTVALIHSVANNKG-VNIDRG-VLK
   :|| :|||:|||| :|| :|||:|:| :| :|||:||||:||||:|:| :||| :||
2  KYEVFRTEEEEEKIKSQGDVTSSVYFMKQTISNACGTIGLIHAIANNKDKMHFESGSTLK
      70      80      90      100     110     120

      120     130     140     150     160     170
1  DFLEKTASLSPEERGRALEKDEKFTADHEALAQEGQTNAAN-HEKVIHHFIALVNKEGTL
   :|||:|:|:|:|:| :|| :| :| :||| :|||:|:| :||| :||| :|
2  KFLEESVSMSPEERARYLENYDAIRVTHETSAHEGQTEAPSIDEKVDLHFIALVHVDGHL
      130     140     150     160     170     180

      180     190     200     210     220
1  YELDGRKSFPIKHGPTSEETFVKDAAKVCKEFMARDPNEVRFTVLALTAQAQQ
   |||||:|:|:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|
2  YELDGRKPFPINHGETSDETLLEDAIEVCKKFMERDPDELRFNAIALSAA
      190     200     210     220     230

```

Figure 5.7B SCORES

Init1: 226 Initn: 423 Opt: 527
44.4% identity in 223 aa overlap

```

      10      20      30      40      50      60
1  MLTWTPLESNPEVLTKYIHKLGVS PAWSVTDVIGLEDDTLEWIPRPVKAFILLFPCSETY
   : :||| :|||:|:| :||| :|| :|||:|:|:|:| :| :| :||| :|
2  MQLKPMEINPEMLNKVLSRLGVAGQWRFDVLGLEEESLGSPAPACALLLFLPLTAQH
      10      20      30      40      50

      70      80      90      100     110
1  EKHRTEHDRIKEVEEQHPEDLFYMRQFTHNACGTVALIHSVANNKG-VNIDRG-VLKDF
   || :||| :| :|| :|| :|||:|:| :| :|||:||||:||||:|:| :||| :||
2  ENFRKKQIEELK--GQEVSPKVYFMKQTIGNSCGTIGLIHAVANNQDKLGFEDGSVLKQF
      60      70      80      90      100     110

      120     130     140     150     160     170
1  LEKTASLSPEERGRALEKDEKFTADHEALAQEGQTNAANHEKVIHHFIALVNKEGTLYEL
   :|||:|:|:|:|:| :|||:| :|:|:|:|:| :| :||| :|| :| :| :|||
2  LSETEKMSPEDRAKCFEKNEAIIQAAHDAVAQEGQCRV--DDKVNHFHIFLNNVDGHLYEL
      120     130     140     150     160     170

      180     190     200     210     220
1  DGRKSFPIKHGPTSEETFVKDAAKVCKEFMARDPNEVRFTVLALTAQAQQ
   |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2  DGRMPFPVNHGASSEDLLKDAAKVCREFTEREQGEVRFSAVALCKAA
      180     190     200     210     220

```

B. DNA Level

The DNA sequence of ovD19C2 was used to search homologous sequences existing in the data base GenEMBL using the FASTA program. The highest score is a cDNA encoding the human neuron specific protein PGP9.5 (Day and Thompson, 1987); these two cDNAs have 57.0% identity. The homology shared by the two cDNA is found in the coding regions of the two cDNAs.

5.5.5 Comparison the ovD19C2 protein structure to the human proteins and yeast protein.

By using the COOH-terminal ethyl ester of ubiquitin, a generic substrate, Wilkinson and his colleagues have identified four proteins showing the ubiquitin carboxyl terminal hydrolase activities from bovine thymus (Wilkinson *et al.* 1986). The four enzymes are all thiol proteases and show a high-affinity binding site for ubiquitin (Mayer and Wilkinson, 1989). The human UCH-L3 gene has been cloned by screening a human B-cell cDNA library using antibodies raised against bovine UCH-L3. Two lines of evidence strongly support that this cDNA does encode the human UCH-L3. First, the fusion protein encoded by this cDNA cross-reacts to the antibodies raise against purified bovine UCH-L3. Secondly, the fusion protein encoded by this cDNA is enzymatically active and is able to hydrolyse the ubiquitin ethyl ester (Wilkinson *et al.* 1989).

From the alignment in Figure 5.8 several distiguishing characteristics of this class of enzymes have been revealed. First, since it has been shown these enzymes are thiol hydrolases, the only conserved Cys residue at position 100 must define the active site (Neurath, 1984). Second, one of the His residues at position 107 and 181 must be essential for catalytic activity (Blow *et al.*, 1969). Finally, since ubiquitin is highly conserved in all eukaryotes, the domain(s) which confer the ubiquitin binding site must also be conserved. The best candidate regions for the ubiquitin binding site are the regions from 43 to 60 and 190 to 198 (Wilkinson *et al.*,

Figure 5.8

An alignment of three ubiquitin carboxyl terminal hydrolases and the predicted ovD19C2 protein.

The amino acid residues identical to *Drosophila* sequence are shaded. The only conserved Cys at position 100, and the two conserved His at position 107 and 177 are outlined.

y-UCH = yeast enzyme

h-UCH-L3 = human ubiquitin C-terminal hydrolase isozyme L3

h-PGP 9.5 = human ubiquitin C-terminal hydrolase isozyme L1

FIGURE 5.8

| | | | |
|-----------|---------------------|---|--------------------------------------|
| | 1 | | 50 |
| y-UCH | MSGENRAVVPIESNP | ELFTNFAHKLGLKNEWAYFDIYSLTEPELLAFLPR | |
| h-PGP 9.5 | | MLNKVLSRLGVAGQWRFDVLGL.EEESLGSVPA | |
| h-UCH-L3 | MEGQ.R.WLPLEANPEVTN | QFLKQGLHPNWQFVDVYGM.DPELLSMVPR | |
| ovD19C2 | | MLTWTPLESNP | EVLTKEYIHKLGVSPAWSVTDVIGL.EDDTLEWIPR |
| | 51 | | 100 |
| y-UCH | PVKAIVLLFPINE..... | DRKSSTSQQITSCYDVIWFKQSVKNAC | |
| h-PGP 9.5 | PACALLLLFP | LTAQHENFRKKQIEELK..GQEVSP..KVYFMKOTIGNSC | |
| h-UCH-L3 | PVCAVLLFP | ITEKYEVFRTEEEEEKIKSQGDVTS..SVYFMKOTISNAC | |
| ovD19C2 | PVKAFILLFPC | SETYEKHRTEEHDRIKEVEEQHPE..DLFYMROFTHNAC | |
| | 101 | | 150 |
| y-UCH | GLYAIILHSLSNNQSL | ..EPGSDLDNFLKSQSDTSSSKNRFDVTTDQFVL | |
| h-PGP 9.5 | GTIGLIH | AVANNQDKLGFEDGSVLKQFL.SETEKMSPEDRAKCFEKNEAI | |
| h-UCH-L3 | GTIGLIH | AIANNKDKMHFESGSTLKKFL.EESVSMSPPEARAYLENYDAI | |
| ovD19C2 | GTVALIHS | VANNKG.VNIDRG.VLKDFL.EKTASLSPEERGRALEKDEKF | |
| | 151 | | 200 |
| y-UCH | NVIKENVQTFSTGQSE | APPEATADTNLFIYITYVEENG | GIFELDGRNLSGPS |
| h-PGP 9.5 | QAAHDAVAQ..EGQC... | RVDDKVNFIHFI | LFNNDGHLYELDGRM....P |
| h-UCH-L3 | RVTHETSAH..EGQTE | APSIDEKVDLFI | FIALVHVDGHLYELDGRK....P |
| ovD19C2 | TADHEALAQ..EGQT | NAAHHEKVIHFI | FIALVNKEGTLYELDGRK....S |
| | 201 | | 245 |
| y-UCH | YLGKSDPTATDLIEQEL | VVRVSYMENANEEDVLNFAMLGIGPNWE | |
| h-PGP 9.5 | FPVNHGASSED | TLLKDAAKVCREFTEREQGE..VRFS | AVALCKAA. |
| h-UCH-L3 | FPINHGETS | DETLLDAIEVCKKFMERDPDE..LRF | NAIALSAA.. |
| ovD19C2 | FPIKHGPTSEET | FKDAAKVCKEFMARDPNE..VRFT | VLALTA |

1989), since the amino acid residues in these two regions are well conserved.

Apart from the high level of identity to the two human ubiquitin hydrolases UCH-L3 and PGP 9.5 (shown in Figure 5.7AB), the putative ovD19C2 protein possesses the important features that define the ubiquitin COOH-terminal hydrolase as described above. From the alignment in Figure 5.8, the Cys residue at the position 100 and the two His residues at positions 107 and 181 are conserved when compared to the two human enzymes and the yeast UCH1. In addition, the putative ubiquitin binding domains at positions 43-60 and 190-198 are also very similar to the human and yeast enzymes. It is therefore likely that OvD19C2 encodes the ubiquitin carboxyl terminal hydrolase of *D. melanogaster*, though final proof of this will require a functional test of the enzymatic activity of this protein.

5.6 Sequencing Analysis of ovD19C1

From the results of in situ hybridization to the polytene chromosomes, the 1.8 kb cDNA ovD19C1 arose from the ligation of two heterologous cDNA molecules, sequencing analysis of this cDNA should be able to provide further evidence supporting this conclusion and provide information about the sequence different from ovD19C2.

Both strands of the region of ovD19C1 which is heterologous to ovD19C2 were sequenced. Only one strand was sequenced for the region homologous to ovD19C2.

The whole sequence of ovD19C1 is presented in Figure 5.9, and it contains 1,809bp nucleotides.

An alignment was made to compare the sequences between ovD19C1, ovD19C2 and the genomic sequence (data not shown). This reveals that ovD19C1 covers the entire length of ovD19C2. In addition, ovD19C1 has 44 bp sequence extending from the 5'end of ovD19C2. This region is homologous to GenD19 and ends at the position 726 nucleotide.

There is a stretch of poly(T) at the presumptive 3'end of ovD19C1. This poly(T) sequence probably represents the poly(A)

Figure 5.9

DNA sequence of ovD19C1.

The sequence homologous to ovD19C2 and GenD19 is referred to as ovD19C1A, the remaining sequence is called ovD19C1B.

Figure 5.9 DNA Sequence of ovD19C1

TTTTTTTTTTT

11 TTTTTTTCGT GTATTTTAC ATTTATTTGT GTAAAGCAAA CGGATTTTTA
61 TCAAGTCTGC GCCATGTATT TGTCTTGGCT ACAGCTGGCT ACCACTGCTT
111 AGTCCTCCTC CTCTTCAGCA GCTCCACCGC CACCGCCCTG GTTCTTGCAG
161 GTTCTTCCTC TTGACGCGAC CGGGACGGCC GCGTCCGAAG GGGCATCTTC
211 AGGGAGAAGT TCGATGTTGC TTCTGGGAGT CCAGGCGTCA CGACGAACGA
261 CGGGATTGTT GACCACCTGC TTGCGAGCAC GAATGTACGC TGACGGATCA
311 GGACGCGAGC ATGATGGATG GACTTGGCAA GTCCAGCTTG AACACCTGCG
361 TCTGCAGACG ACGCTCCAAG AAGTCCTGAA TCTTCAGACC CAGCACGTAA
411 TCGAGCTTCA TGCGGACTCT TCAGGACACC GATACGGACC AGACGGCGCA
461 CGAGGGACTT ACCCTGGAAC AGACGCTTCG CGTCCTTCTC GTCGAGGGTC
511 AGCAGCTCAC GAGCGACCTT ACGGATCTTA GCCAGGGCGT ACTTGACCGC
561 CACACTTCGC GCTTGTTGCA GACCATACTC GCCGATGATC TTCAACTCCT
611 GGTCCAGACG CGCCTTCTCA TAGGGGCGAC GGGGAGTCAC GTAGGTCTTC
661 GAGAAGACCG AGGGTATGCG GCCGTTTACC ATGGTTCCGA CAATTCAACC
ovd19C1B-->ovD19C1A
711 CAGCAGGTTT TGCTCACATT CAGAAAAAAG TAAAGCTCGT CAAAAGTGCC
761 GGTAATTGTG TGTAGAAATC CTTCATTTTC CTGTGAATTT TCAATCAGAT
811 CACTGCTAAA GCCAAATTCA AAAATATGTT AACCTGGACG CCACTTGAAT
861 CTAATCCCGA GGTTTTGACC AAGTACATAC ATAAACTGGG CGTGTGCGCA
911 GCCTGGTCGG TAACTGACGT CATTGGATTG GAAGATGACA CCTTGGAATG
961 GATTCCGCGT CCCGTAAAGG CGTTCATTTT GCTCTTCCCG TGCAGCGAAA
1011 CCTATGAGAA GCACCGCACC GAGGAGCACG ATCGGATTAA GGAGGTGGAG
1061 GAGCAGCATC CCGAGGATCT CTTCTACATG CGCCAGTTCA CCCACAACGC
1111 CTGTGGAACC GTCGCCCTGA TCCACAGCGT GGCCAACAAC AAAGGAGTGG
1161 ACATTGACCG CGGAGTACTG AAGGACTTCC TGGAGAAGAC AGCTTCCTTG
1211 TCCCCGGAGG AACGCGGACG GGCCCTCGAA AAAGACGAGA AGTTCACCGC
1261 CGATCATGAG GCCTTGCTC AAGAGGGCCG ACGAATGCCG CCAATCATGA
1311 GAAGGTGATC CACCACTTCA TCGCACTGGT GAACAAGGAG GGTACTCTGT

Figure 5.9 continued

1361 ACGAGCTGGA TGGCCGCAAG TCCTTCCCGA TCAAGCACGG ACCGACTTCG
1411 GAGGAGACTT TTGTGAAGGA TGCGGCCAAG GTCTGCAAGG AGTTCATGGC
1461 TCGCGATCCC AACGAAGTGC GCTTCACCGT TTTGGCCTTG ACCGCGGCAC
1511 AACAATAGGG TGGA ACTAAT CCCATAACTG AGAACCACCA AGATCCCGGA
1561 TCTGCCAGCT ATATATTTAT ATATCCAACA TATTTTAAAC CCAGGCTGCA
1611 AAGGCGAACG AAGCCCGCCT TTCAGCCGGT CAATAGTTTA CGCATATTTA
1661 TTCCGCATTG ACCCAAAGCC CCATATGTAT TGGGAACACC ATGATTGTTA
1711 TTAAAGAGTG CACTCTAGCT TGATGCTAAG CTTATACTGA AAAAAAAAAA
1761 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Figure 5.10 DNA sequence of ovD19C1B and predicted amino acid sequence

5' TTACCGGCACTTTTGACGAGCTTTACTTTTTTCTGAATGTGAGCAAAACCTGCTGGGTTG
AATTGTCGGAACCATGGTGAACGGCCGCATACCCTCGGTCTTCTCGAAGACCTACGTGACT
MetValAsnGlyArgIleProSerValPheSerLysThrTyrValThr-
CCCCGTCGCCCTATGAGAAGGCGCGTCTGGACCAGGAGTTGAAGATCATCGGCGAGTAT
ProArgArgProTyrGluLysAlaArgLeuAspGlnGluLeuLysIleIleGlyGluTyr-
GGTCTGCAACAAGCGCGAAGTGTGGCGGTCAAGTACGCCCTGGCTAAGATCCGTAAGGTC
GlyLeuGlnGlnAlaArgSerValAlaValLysTyrAlaLeuAlaLysIleArgLysVal-
GCTCGTGAGCTGCTGACCCTCGACGAGAAGGACGCGAAGCGTCTGTTCCAGGGTAAGTCC
AlaArgGluLeuLeuThrLeuAspGluLysAspAlaLysArgLeuPheGlnGlyLysSer-
CTCGTGCGCCGTCTGGTCCGTATCGGTGTCTCTGAAGAGTCCGCATGAAGCTCGATTACGT
LeuValArgArgLeuValArgIleGlyValLeuLysSerProHisGluAlaArgLeuArg-
GCTGGGTCTGAAGATTCAGGACTTCTTGGAGCGTCGTCTGCAGACGCAGGTGTTCAAGCT
AlaGlySerGluAspSerGlyLeuLeuGlyAlaSerSerAlaAspAlaGlyValGlnAla-
GGAATTGCCAAGTCCATCCATCATGCTCGCGTCCTGATCCGTCAGCGTACATTCGTGCTC
GlyLeuAlaLysSerIleHisHisAlaArgValLeuIleArgGlnArgThrPheValLeu-
GCAAGCAGGTGGTCAACAATCCCGTCGTTTCGTGACGCCTGGACTCCCAGAAGCAACA
AlaSerArgTrpSerThrIleProSerPheValValThrProGlyLeuProGluAlaThr-
TCGAACTTCTCCCTGAAGATGCCCCTTCGGACGCGGCCGTCCCGGTTCGCGTCAAGAGGAA
SerAsnPheSerLeuLysMetProLeuArgThrArgProSerArgSerArgGlnGluGlu-
GAACCTGCAAGAACCAGGGCGGTGGCGGTGGAGCTGCTGAAGAGGAGGAGGACTAAGCAG
GluProAlaArgThrArgAlaValAlaValGluLeuLeuLysArgArgArgThrLysGln-
TGGTAGCCAGCTGTAGCCAAGACAAATACATGGCGCAGACTTGATAAAAAATCCGTTTGCT
Trp
TTACACAAATAAATGTAAAAATACACGAAAAAAAAAAAAAAAAAAAA-3'

Figure 5.11

An alignment between ovD19C1B protein and the ribosome protein 1024 of Dictyostelium discoideum.

The identical amino acid residue in the two proteins are shaded.

Dr124 = amino acid sequence from ovD19C1B

R124\$Dicdi = amino acid sequence of the ribosome protein 1024 of D. dicodum

FIGURE 5.11

| | | |
|-------------|---|-----|
| | 1 | |
| Dr124 | MVNGRIPSVFSKTYVTPRRPYEKARLDQELKIIIGEYGLQQARSV.AVKYA | |
| R124\$Dicdi |MSSNYSKTSHTPRRPFEKERIDAELKVVGEFGLKNKNEVWRVQYA | |
| | 51 | 100 |
| Dr124 | LAKIRKVARELLTLDEKDAKRLFOGKSLVRRRLVRIGVL.KSPHEARLRAG | |
| R124\$Dicdi | LAKIRKAARELLVLDEKDPKRIFEGSALLRRLHKLGVMEESKNKLDYILN | |
| | 101 | 150 |
| Dr124 | SEDSGLLGASSADAGVQAGLAKSIHHARVLIRQRTFVLASRWSTIPSEFV | |
| R124\$Dicdi | LKVQDFMERRLQTLVFKNGLAKSIHHARVLIKGRHIRVGKQLVNVPSFLV | |
| | 151 | 200 |
| Dr124 |TPGLPEATSNFSLKMPLRTRPSRSRQEEEPARTRAVAVEL | |
| R124\$Dicdi | RVESQKHLGLASTSPLAGGRPGRRARKMAKNNSKGEEN..... | |
| | 201 | 230 |
| Dr124 | LKRRRTKQW | |
| R124\$Dicdi | | |

tail of the other cDNA which was ligated to the cDNA sequence represented by ovD19C2. When the sequence of ovD19C1 is reversed, the poly(T) "head" becomes a poly(A) tail headed by a poly(A) addition signal AAATAAA 14 bp upstream. This result confirms the conclusion that ovD19C1 is a ligation of two cDNAs.

The sequence of ovD19C1 is thus divided into ovD19C1A and ovD19C1B; ovD19C1A represents the sequence homologous to GenD19, while ovD19C1B represents the remaining 715 bp sequence of ovD19C1.

When the reversed sequence of ovD19C1B is translated into an amino acid sequence, the longest ORF encodes a polypeptide of 180 amino acids (Figure 5.10). The translation initiation codon ATG is located at position 74 headed by a AACC sequence consensus which matches the translation initiation sequence consensus of *Drosophila* genes (Cavener, 1987). The codon usage in this ORF is similar to many *Drosophila* genes. In respect of encoding region, ovD19C1B is therefore considered as full length.

The inferred protein was used to search the protein data base Swissprot. The best match is the ribosomal protein 1024 of *Dictyostelium discoideum* (Steel *et al.*, 1987). The *Drosophila* and *D. discoideum* proteins have 61 identical amino acid residues and 67 conserved substitutions. Most of the homologous sequences are in the C-terminal region (Figure 5.11).

5.7 Discussion

By screening a *Drosophila* ovarian cDNA library using the genomic fragment GH4, two cDNA clones (ovD19C1 and ovD19C2) have been isolated. This genomic fragment is located 7 kb upstream of the P-transposon P[lac, ry⁺], and it detects a 1.1 kb RNA transcript which shows a similar expression pattern to the *lacZ* fusion gene in the strain D19. Though it should be noted that since there is another transcription unit between the p[lac, ry⁺] insertion and the gene we have studied may not actually be the region of the genome responsible for the *lacZ* expression pattern. OvD19C1 is 1.8 kb in length, and this cDNA only detects a 1.3 kb HindIII genomic fragment which is situated in the middle of the genomic

clone λ D10 (Figure 4.4). This fragment has been named GH4. In situ hybridisation to the polytene chromosomes shows that this cDNA hybridises to two chromosomal locations; one is at 98F on the right arm of chromosome 3 where P[lac, ry⁺] maps, and the other is on the left arm of chromosome 3 (Figure 5.3C). This suggests that ovD19C1 represents the transcripts of two different genes. The comparison between ovD19C1 and ovD19C2 shows that ovD19C1 covers the entire sequence of ovD19C2. Since ovD19C2 only detected one signal on polytene chromosomes (Figure 5.3B), the signal on chromosome 2 detected by the ovD19C1 probe must represent another gene which does not contain sequences homologous to ovD19C2. There are poly(A) tails at both ends of ovD19C1, and both are headed by a polyadenylation signal. From these lines of evidence, ovD19C1 appears to be a ligation of two cDNA molecules resulting from cloning artefact. This probably happened when the EcoRI linker was added to the cDNA molecules during the construction of the cDNA library.

OvD19C2 is a 1 kb cDNA and represents the 1.1 kb transcript detected by the genomic fragment GH4 (Figure 5.1). Southern analysis shows that the ovD19C2 probe hybridises to a 1.3 kb HindIII fragment, and suggests that ovD19C2 is a single copy gene. OvD19C2 maps to 98F on the third chromosome where P[lac, ry⁺] maps.

Sequencing analysis shows that GH4 covers the whole sequence of ovD19C2. Although the cap site of ovD19C2 has not yet been determined, a TATA box is found at the position 5x bp upstream of the 5'end of this cDNA. Comparison of the two sequences has revealed that this gene contains 3 introns. The intron recognition sequences are typical of the *Drosophila* consensus for intron splicing (Keller and Noon, 1985).

The longest open reading frame from ovD19C2 encodes a 227 amino acid protein with a molecular weight of 26 kD. Several lines of evidence suggest that this ORF may be the coding frame of this gene. The translation initiation sequence matches the *Drosophila* consensus sequence very well (Cavener, 1987). The codon usage in

the ORF is typical of *Drosophila* genes (Grantham *et al.*, 1980), and the longest ORF in the two other frames only encode peptides of 72 and 63 amino acids.

Data base searches for genes homologous for DNA and protein were conducted. When the deduced protein from ovD19C2 was used to scan the protein data base Swissprot, it showed high similarity (51.1%) to human UCH-L3 in the protein data base Swissprot. The two proteins have 115 identical amino acid residues and 77 conserved changes. The ovD19C2 protein also shows high similarity to the human neuron protein PGP 9.5 in the data base Prosrch (Collins and Coulson, unpublished). The score (754) for similarity between the two sequences is very high, which means that the similarity between the two proteins is unlikely to be caused by random pairing (Coulson, personal communication). The PGP 9.5 protein is known to be the ubiquitin COOH-terminal hydrolase iszyme L1 (Wilkinson *et al.*, 1989)

Various lines of evidence suggest that the *Drosophila* ovD19C2 protein is a ubiquitin carboxyl terminal hydrolase. Firstly, the predicted protein has a molecular mass about 26 kD: the same as those of human UCH-L3 and the yeast YUH1 (Wilkinson *et al.*, 1989; Miller *et al.*, 1989). Secondly, it shares high level of similarity to the two human enzymes UCH-L3 and PGP 9.5. Furthermore, ovD19C2 protein is conserved at various regions in the two human protein and the yeast proteins, which define this family of enzymes (Wilkinson *et al.*, 1989). Since ubiquitin is highly conserved in all eukaryotic organisms (see following chapter), the ubiquitin hydrolases probably contain of conserved sequence which binds to ubiquitin. The putative ubiquitin binding domains include the amino acid residues 43-60 and 190-198 in the alignment shown in Figure 5.8 (Wilkinson *et al.*, 1989). It can seen from this figure that ovD19C2 protein is also conserved in these two regions. Mayer *et al.* (1989) have demonstrated that UCH-L3 is a thiol protease, and that one of its cysteines will define the active site of the enzyme (Neurath, 1984). In the alignment of the three enzymes, the only conserved Cys residue is the one at position 100, and this Cys is also found in the ovD19C2 protein. There are two His residues at

positions 107 and 181, which are conserved among the four proteins. It is known that His is catalytically essential for proteases (Blow *et al*, 1969).

Sequencing analysis of ovD19C1B indicates that this cDNA encodes a protein of 190 amino acids. This predicted protein shares considerable similarity to the ribosome protein 1024 of *Dictyostelium* (Steel *et al*, 1969). The similarity between the two proteins is very high among the first 150 amino acid residues at their C-termini with several significant long matches. This suggests that the two proteins may have similar function.

CHAPTER SIX

ANALYSIS OF THE EXPRESSION PATTERN OF OVD19C2 AND OVD19C1B

6.1 Introduction

Acquiring its name by its presence in all eukaryotic organisms, ubiquitin is the most conserved protein in eukaryotes. Ubiquitin contains 76 amino acid residues; its amino acid sequence is identical in mammals, and it only differs by three amino acids from yeast to human (Ozkaynak *et al.*, 1987). Ubiquitin can be identified in its free form or as covalently attached to proteins in various cellular compartments (reviewed by Finley and Chau, 1991).

Ubiquitin is encoded by a family of fusion genes (reviewed by Schlesinger and Bond, 1987). This gene family comprises of two classes. The first class consists of genes with tandem repeats, which encode a head to tail polyubiquitin fusion protein. The final repeat of the COOH-terminal of the polyubiquitin fusion is followed by a non-ubiquitin amino acid. The copy number and the number of the ubiquitin repeats vary between different species. Polyubiquitin genes have been identified in yeast (Ozkaynak *et al.*, 1987), *Dictyostelium* (Ohmachi *et al.*, 1989), plants (Gausing and Barkardottir, 1986), *Xenopus* (Dworkin-Rastl *et al.*, 1984), chicken (Bond and Schlesinger, 1985) and humans (Wiborg *et al.*, 1985); and have also been identified in *D. melanogaster* (Arribas *et al.*, 1986; Lee *et al.*, 1988). There is only one copy of the polyubiquitin gene in *Drosophila*, consisting of 18 ubiquitin repeats. The second class of ubiquitin genes encode a single ubiquitin and an in frame COOH-terminal protein extension (CEP). Two types of the ubiquitin carboxyl terminal extension protein genes (UbCEP gene) have been identified. The first encodes ubiquitin and a CEP consisting of 52 amino acids (CEP52); and in the second type, the CEP contains 76-80 amino-acids (CEP76-80). The CEPs are also highly conserved at the amino acid level among very divergent organisms, from yeasts to humans. In *Drosophila*, only a UbCEP80 gene has been identified (Lee, *et al.*, 1988) with its CEP showing 65% similarity to the yeast sequences and 82% to human sequences. Inspection of the CEPs shows that they all contain a common nucleic acid binding motif known as a "zinc finger".

Expression patterns of the two classes of ubiquitin genes show differential regulation. In *Dictyostelium*, the polyubiquitin gene is

expressed throughout most developmental stages, but is switched off during the growth stage in which the expression of the UbCEP genes is predominant (Ohmachi *et al.*, 1989). In yeast, the polyubiquitin gene is expressed at low levels during the exponential growth phase. The expression is dramatically increased as the cells enter the stationary growth phase. As in *Dictyostelium*, the CEPs are only expressed in rapid dividing yeast cells. A similar expression pattern was also observed in barley; and the polyubiquitin transcript is constitutively expressed in vegetative tissues while the UbCEP transcript is only detected in dividing cells (Gausing and Barkardottir, 1986). In *D. melanogaster*, the polyubiquitin gene encodes a constitutive mRNA (Lee *et al.*, 1988).

In most eukaryotes, the expression of the polyubiquitin gene can be induced under stress, such as heat shock and starvation. Heat shock regulatory elements have been identified in the 5' untranslated region of the polyubiquitin genes in yeast (Ozkaynak, *et al.*, 1987), *T. cruzi* (Swindle, *et al.*, 1988) and chicken (Bond and Schlesinger, 1985). In *D. melanogaster*, although a heat shock consensus has not been identified, the transcription of the polyubiquitin gene was induced three fold by heat shock (Lee, *et al.*, 1988), but in contrast, the UbCEP genes are shut down during stress. In *D. melanogaster*, the UBCEP80 transcript is detected at normal levels under heat shock (Lee *et al.*, 1988).

The covalent attachment of ubiquitin to a number of cellular proteins has been demonstrated to be involved in a variety of different cellular process, such as selective protein degradation, chromatin structure, stress response and ribosome biogenesis (reviewed by Finley and Chau, 1991). It is estimated that 90% of short-lived proteins are subjected to degradation via the ubiquitin-mediated proteolysis (Finley *et al.*, 1984). Ubiquitin is found covalently attached by an isopeptide bond to the epsilon amino group of lysine of the target protein through its C-terminus. This is joined by an enzyme complex which transfers ubiquitin to the target protein. In this process, ubiquitin is first attached to the E1 enzyme which activates ubiquitin, then, ubiquitin is transferred to another enzyme E2. Ubiquitin is finally transferred from the E2 to

the target proteins to form an isopeptide bond-linked conjugates. Once ubiquitin is attached, the target protein is then digested by a ubiquitin dependent, ATP requiring protease complex, and ubiquitin is released (Picart and Rose 1985a; Haas and Bright, 1988).

Histone proteins are found to be the most abundant ubiquitinated protein in cells. The attachment of ubiquitin leads to histone proteins undergoing degradation. The function of histone ubiquitination is not fully understood. Limited information suggests that ubiquitination may have a function to maintain the chromatin in a transcriptionally active state. In *Drosophila* and other organisms, the ubiquitinated H2A and H2B histones are preferentially present in a transcriptionally active region (Levinger and Varshavsky, 1982; Nickel *et al.*, 1989). Ubiquitin is removed from histones during chromosome condensation (Mueller, *et al.*, 1985). Ubiquitination and deubiquitination are reversible processes; yet genes active in transcription are not always rich in ubiquitinated histones (Huang, *et al.*, 1986).

As stated earlier, the polyubiquitin gene is inducible during conditions of stress. In *S. cerevisiae*, deletion of the polyubiquitin gene UBI4 leads to cells hypersensitive to various conditions of stress, such as heat shock, and starvation (Finley, *et al.*, 1987). The function of UBI4 during stress is to produce mono-ubiquitin, and it is proposed that ubiquitin is required to target abnormal proteins induced during stress conditions.

Another discovery about the function of ubiquitin genes is the participation of UbCEPs in ribosome biogenesis. In *S. cerevisiae*, the UbCEP gene products are rapidly processed to free ubiquitin and the carboxyl extension proteins during or immediately after translation. The two CEP52 proteins expressed from UBI1 and UBI2 are assembled in the large subunit of the ribosome, while UbCEP72 is produced by UBI3 associated with the ribosomal protein S34 (Finly *et al.*, 1989; Reelman and Rechsteiner, 1989). Deletions of any of the three genes can lead to a slow growth phenotype; and double mutation of the two CEP52 loci (*ubi1Δ/ubi2Δ*) is cell-lethal. Analysis of these deletions indicates that the slow growth

phenotype is caused by ribosome deficiencies (Finley *et al.*, 1989). Although ubiquitin itself is not incorporated in the ribosome, expressing CEP72 without the N-terminal fused ubiquitin inefficiently complements the *ubi3Δ* phenotype (Finley *et al.*, 1989). This suggests that the fusion to ubiquitin significantly enhances the function of CEP72. CEPs have also been identified associated with ribosomal proteins in human cells (Monia *et al.*, 1989). When the human UbCEP genes are expressed in *S. cerevisiae* cells, the gene products are also rapidly cleaved to free ubiquitin and CEP tails. In addition, expression of human UbCEP80 fully complements the yeast *ubi3Δ* phenotype (Monia *et al.*, 1989). The ability of CEPs to substitute each other between such diverged organisms suggests that the function of CEPs is highly conserved, and possibly so do the enzymes supporting this function. Like the expression of yeast CEP72 without the N-terminal ubiquitin could not rescue the *ubi3Δ* phenotype, expression of human CEP80 without the N-terminal ubiquitin, or expression of human CEP80 and ubiquitin in separate plasmids, failed to rescue the *ubi3Δ* phenotype. Analysis of expression levels of mRNA and proteins reveals that although the transcription of both UbCEP80 and CEP80 is at the same level, the protein translated from the UbCEP80 mRNA is 5-10 times higher than that from CEP80 mRNA (Finley *et al.*, 1989). This indicates that ubiquitin is required for efficient expression of CEPs.

The finding that CEPs are ribosomal proteins may help to explain the reason for reciprocally-coordinated regulation of their genes with polyubiquitin genes. In rapidly growing cells, protein synthesis is predominant: large amount of ribosomes are required. In this condition, the expression of UbCEP genes is greatly enhanced. By contrast, during conditions of stress, the priorities of cells are to target and degrade abnormal proteins. The expression of polyubiquitin genes becomes maximal while the expression of UbCEP genes is switched off.

In all the cellular processes discussed where ubiquitin is involved, enzymatic activities for deubiquitination are required. In the ubiquitin mediated protein degradation, deubiquitinating enzymes are required to regenerate functional ubiquitin from E1-

ubiquitin and E2-ubiquitin conjugates before ubiquitin is finally attached to the target protein. In addition, a protease-isopeptidase complex can start at one end of the ubiquitinated protein, and release free ubiquitin in the process of cleaving the peptide chain. Isopeptidase activities are expected to process the isopeptide bond between ubiquitin and the attached protein which does not undergo proteolysis, such as histones. A third class of enzyme is involved in the post-translational processing of the products from ubiquitin encoding genes. All the ubiquitin genes encode a hybrid protein where ubiquitin is either fused to itself as polyubiquitin, or to a heterologous protein extension through its C-terminus as UbCEPs. In the conditions of stress, polyubiquitin is rapidly processed to mono-ubiquitin which targets to abnormal proteins. Ubiquitin is removed from UbCEPs before CEPs join the ribosome complex.

Deubiquitinating these diverse conjugates is performed by a set of enzymes. Four genes encoding enzymes, UBP1-3 and YUH1, have been cloned in yeast which cleave ubiquitin-protein fusions (Tobias *et al.*, 1991, Miller *et al.*, 1989). UBP1-3 are within the same enzyme family which shows little sequence similarity to YUH1. It has been shown that UBP1 can cleave, *in vitro*, yeast ubiquitin fusion proteins UBI2 (CEP52) and UBI3 (CEP72) and the ubiquitin- β -galactosidase fusion. YCH1 has been tested for processing a variety of artificial ubiquitin fusion proteins. It is able to cleave a ubiquitin fusion protein containing a 72 amino acid C-terminal extension (Miller *et al.*, 1989). Five ubiquitin conjugate processing enzymes have been identified in mammalian cells (Mayer and Wilkinson, 1989; Jonnalagadda *et al.*, 1989). Among them are the human PGP 9.5 protein and the human ubiquitin carboxyl terminal hydrolase described in chapter 5. It is to these proteins that the predicted ovD19C2 protein shows high sequence similarity (5.4.2). The human UCH-L3 was only tested for processing small ubiquitin adducts such as ubiquitin C-terminal thiol ester. Its homology to YUH1 implies that it may have the ability to cleave larger ubiquitin fusion proteins such as UbCEPs.

In the preliminary experiments described in chapter 4, the Northern analysis was not comprehensive in terms of specificity of

developmental stages and tissue types. In this chapter, further analysis was carried out to draw a finer picture of the expression pattern during development of ovD19C2 and ovD19C1B.

6.2 Analysis of the Developmental Profile of Ov19C2

6.2.1 Northern Analysis

Northern blots were made with RNA extracted separately from embryos aged 0-2h., 2-4h., 4-6h., 6-8h., 3rd instar male larvae, female larvae, pupae, testes, ovaries; carcasses of adult males and carcasses of adult females. The Northern blots were then probed with ov19C2. After autoradiography, the probe was stripped off and the filters were re-probed with α -tubulin as a loading check. The results are shown in Figure 6.1. It can be seen that a 1.1 kb transcript is detected in almost every tissue type and developmental stages, except the pupal stage. However, the abundance of the mRNA varies greatly between different tissue types and developmental stage. In adults, the 1.1 kb transcript is highly enriched in ovaries and testes (Figure 6.1, lane 9 and lane 8). It is also present at low levels in both male and female carcasses (Figure 6.1, lane 10 and lane 11). In embryos, the 1.1 kb is abundant at 0-2 hour and 2-4 hour embryos, and is detected at decreasingly lower levels in 4-6 and 6-8 hour embryos (Figure 6.1, lane 2 to lane 5). The abundance of this transcript decreases in 4-6 hour-old embryos. The enrichment of this transcript in ovaries and early embryos presumably reflects the expression of a maternally expressed gene transcribed in the ovary and delivering its message to the oocyte, to be used in early embryogenesis.

Another transcript of 1.3 kb in length was detected in 2-4 hour-old embryos with the ovD19C2 probe (figure 6.1, lane 3). As the ovD19C2 probe only detected a single 1.1 kb transcript in the ovary, it is unlikely that this 1.3 kb message is maternally transcribed. Since at this stage (2-4 hour embryos) the zygotic genes are already activated, it may be that this RNA is zygotically derived.

The lower level of the 1.1 kb mRNA detected in larvae and adult carcasses indicates that the activity of ovD19C2 is also required in different tissues and at different developmental stages.

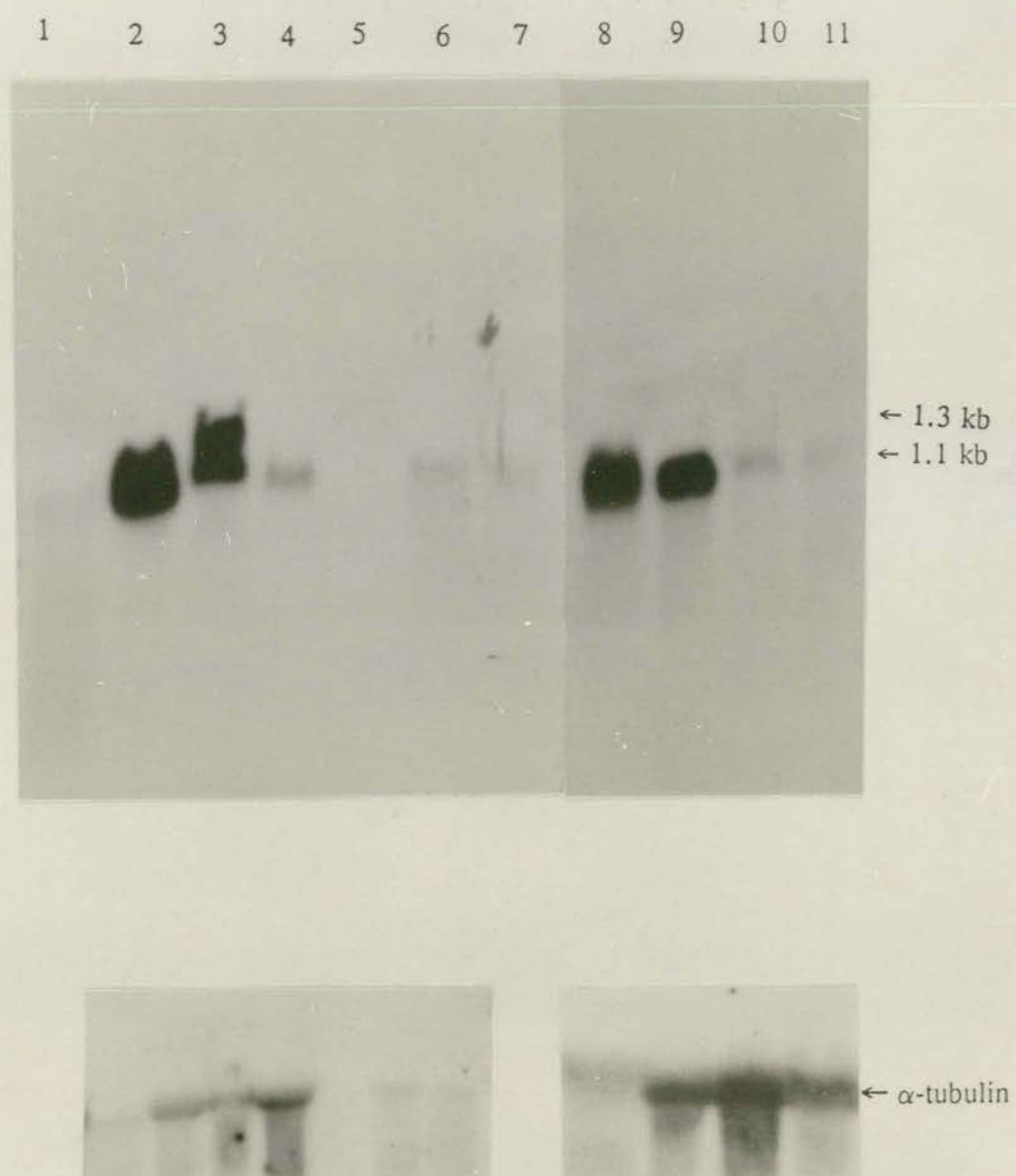
Figure 6.1

Northern blots probed with ovD19C2 probe, showing the developmental profile of this ovD19C2.

- 1 = total RNA from pupae**
- 2 = total RNA from 0-2 hour embryos**
- 3 = total RNA from 2-4 hour embryos**
- 4 = total RNA from 4-6 hour embryos**
- 5 = total RNA from 6-8 hour embryos**
- 6 = total RNA from male 3rd larvae**
- 7 = total RNA from female 3rd larvae**
- 8 = total RNA from testes**
- 9 = total RNA from ovaries**
- 10 = total RNA from adult male carcasses**
- 11 = total RNA from adult female carcasses**

The lower panels show the same filters probed with α -tubulin as a loading check. In lane 5, the RNA from 6-8 hour embryos was obviously underloaded.

FIGURE 6.1



6.2.2 In Situ Detection of the Transcript of OvD19C2

In situ hybridization has been widely used to determine the temporal and spatial expression patterns of developmentally regulated genes such as chorion genes, bicoid, vasa). These techniques usually involve tissue sectioning and using radioactive labelled probes. However, this is laborious and time-consuming. Tautz and Pfeifle (1989) have developed a method using digoxigenin-labelling and subsequent antibody detection of the labelled DNA probe, to detect mRNA in whole mount embryos. This method has greatly simplified procedures, and has been widely used in studies of developmentally regulated genes in *Drosophila*, and in other animals. In this section, I will present the results of determining the spatial distribution of ovD19C2 transcript by in situ hybridization using the method of Tautz and Pfeifle (1989), modified for whole mounted ovaries and testes.

DNA of ovD19C2 was labelled with digoxigenin-11-dUTP as described in 2.5.9.2. For the control, pBluescript DNA was used. Tissue preparation is also described in 2.5.15. except that the intestine was used as a control.

In figure 6.2A, the nurse cells start to express ovD19C2 mRNA at stage 9, and the strongest signal is detected in the nurse cells in stage 10 egg chambers. The signal is still present in nurse cells in stage 11-14 egg chambers (Figure 6.3A). The ovD19C2 transcript is also detected in stage 10 oocytes (Figure 6.2B). A very strong signal is detected in stage 14 oocyte after the chorion and vitelline membrane have been physically removed (Figure 6.3B). The mRNA is homogenously distributed within the ooplasm. This suggests that transcription of the ovD19C2 gene starts at stage 9, and the transcript is later transported into the oocyte, since such a strong signal is unlikely to be generated by transcription in the oocyte nucleus.

Eggs were collected during a period of 16 hours and their developmental stages were determined according to their morphology, as described by Bownes (1982), thus the corresponding age can be estimated. In the embryos, the ovD19C2 transcript is detected from stage 1 to about stage 8. The abundance

Figure 6.2

Detection in situ of mRNA in the ovary using ovD19C2 probe.

A. The hybridization signal is detected in nurse cells of stage 9-11 egg chambers. Scale bar = 100 μ m

B. The hybridization signal is detected in nurse cell and oocyte of stage 10 egg chambers. Scale bar = 50 μ m

O = oocyte

N = nurse cell

FIGURE 6.2

A



B



Figure 6.3

Detection in situ of mRNA in the ovary using ovD19C2 probe.

- A. The hybridization signal is detected in degenerating nurse cells (arrows) of stage 13 egg chambers. Scale bar = 50 μ m
- B. The hybridization signal is detected in oocyte of stage 13-14 egg chambers. Scale bar = 50 μ m
- C. The hybridization signal is detected in the embryo of blastoderm stage. Scale bar = 50 μ m

FIGURE 6.3



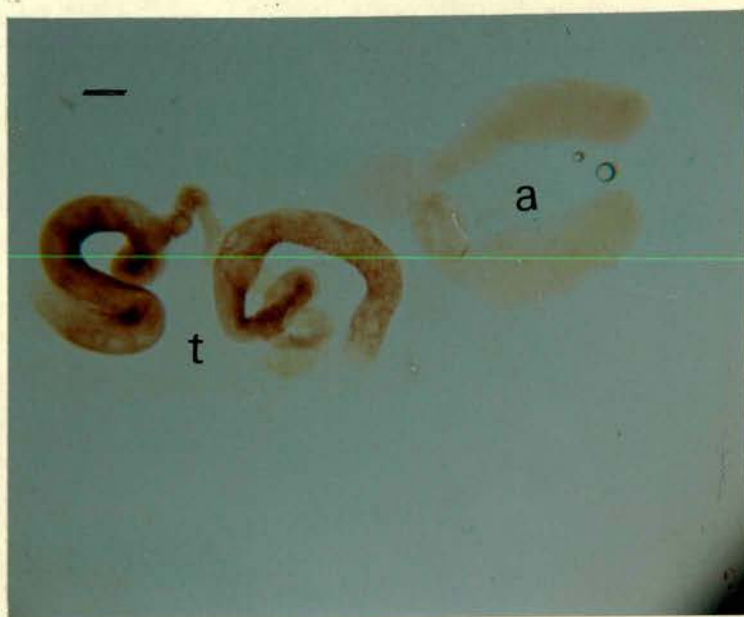
Figure 6.4

Detection in situ of mRNA in the testes using ovD19C2 probe.

A. The hybridization signal is detected in testes (t). The accessory gonads (a) only stained faintly. Scale bar = 100 μm

B. The hybridization signal is detected in testes at the position where spermatocyte cysts are normally situated (arrows). Scale bar = 50 μm

FIGURE 6.4



A



B

of the mRNA decreases as the age of the embryo increases. The strongest signal was detected in stage 1-7 embryos (Figure 6.3C), which corresponds to approximately 0-4 hours development time at 25°C.

In male gonads, the ovD19C2 probe only produces a signal in the testis (figure 6.4A.B). Based on the position where the probe stains, the mRNA of ovD19C2 is apparently transcribed in the spermatocyte. The accessory glands and other gonadal tissue stained faintly.

The expression pattern of ovD19C2 determined by in situ hybridization is in agreement with the previous Northern analysis, as described in 6.1.1.

6.3 Northern Analysis Using the ovD19C1 Probe

6.3.1 Comparison of Transcripts Detected by the OvD19C1 and OvD19C2 Probes

Northern blot analysis was used to compare the difference between ovD19C1 and ovD19C2. Figure 6.5A shows that on the Northern blot, both ovD19C1 and ovD19C2 detected the 1.1 kb transcript previously identified using ovD19C2 (Figure 6.1). From this data, both ovD19C1 and ovD19C2 contain sequences encoding the 1.1 kb message. In addition to this 1.1 kb transcript, ovD19C1 detected another transcript approximately 0.8 kb in length (Figure 6.5A, lane 2). From sequencing data (5.7), the 0.8 kb transcript is apparently encoded by ovD19C1B.

6.3.2 Expression Pattern of OvD19C1 During Development

A Northern blot containing RNA extracted from ovaries, testes, male carcasses, female carcasses, third instar larvae and 0-8 hour-old eggs was probed with ovD19C1. The ovD19C1 probe detected two transcripts in all tissues types tested (Figure 6.5B). The expression pattern of ovD19C1A is similar to that of ovD19C2 as described in 6.1. The lower level expression of the 1.1 kb transcript in the testes RNA track is probably due to under-loading of this sample, as shown using the Adh probe (Figure 6.5B). The 0.8 kb transcript detected by ovD19C1 probe was instead much more

Figure 6.5

A. Northern blots containing total ovary RNA probed with ovD19C1 and ovD19C2 respectively.

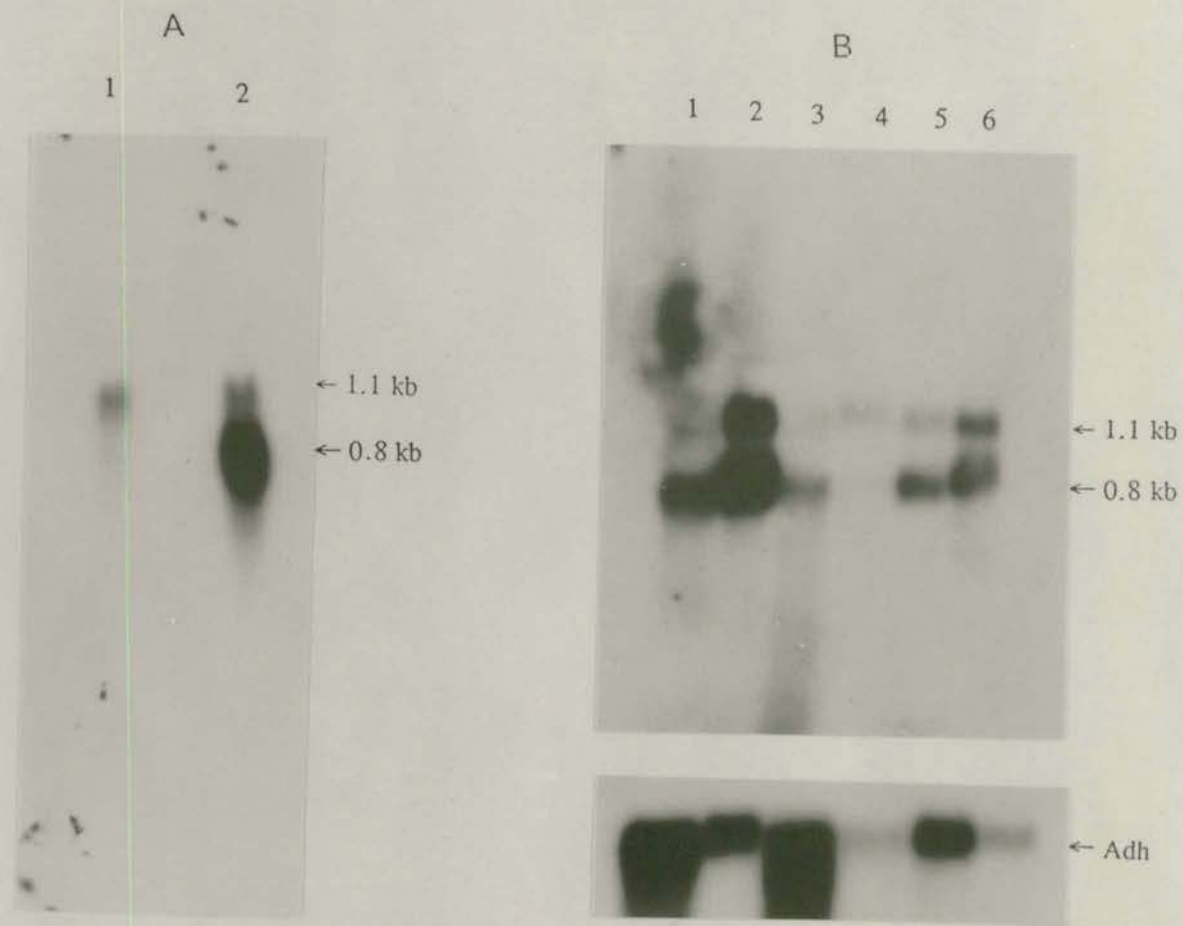
- 1 = ovD19C2 probe
- 2 = ovD19C1 probe

B. Northern blot probed with ovD19C1 probe, showing the developmental profile of this ovD19C1.

- 1 = total RNA from adult female carcasses
- 2 = total RNA from ovaries
- 3 = total RNA from adult male carcasses
- 4 = total RNA from testes
- 5 = total RNA from 3rd instar larvae
- 6 = total RNA from 0-8 hour embryos

The lower panels show the same filters probed with Adh probe as a loading check. In lane 4 and lane 6, the RNA from testes and 0-8 hour embryos were obviously underloaded

FIGURE 6.5



ovarian enriched than the 1.1 kb transcript (Figure 6.5B, Lane 2). Since the testis RNA was underloaded, it is difficult to judge the transcription level of the 0.8 kb mRNA in testes. From Figure 6.5B, it can be estimated that the abundance of this mRNA in the testis is lower than that of the 1.1 kb transcript. From this point of view, the expression of ovD19C1B is enriched in the ovary to a greater extent than ovD19C1A. The 0.8 kb transcript was also detected at a high level in embryos (Figure 6.5B). As the 0.8 kb transcript is enriched in the ovary and embryo, it indicates that this message is the product of a maternally-expressed gene stored in the oocyte. However, some zygotic expression could not be ruled out, since the RNA was prepared from 0-8 hour old eggs.

6.4 Discussion

The expression patterns of the cDNAs ovD19C2 and oVD19C1B were analyzed using Northern blots and in situ hybridization.

Northern analysis shows that ovD19C2 probe detects a 1.1 kb transcript, enriched in ovaries, early embryos and testes. This transcript was detected at a lower level in all the other tissues and developmental stages tested (Figure 6.1). The abundance of this transcript in embryos decreased as the age of the embryo grew. In Figure 6.1, the RNA from 6-8 hour embryos was underloaded, but the decreasing level of this mRNA can also be seen in the 4-6 hour embryos. Since there is rarely transcription within oocytes and early embryos (Edgar and Schubiger, 1986), the high level of ovD19C2 message detected in the early embryos must be derived maternally. This pattern of transcription has been seen within many maternal effect genes. The low level and ubiquitous presence of mRNA in all types of tissues, and during different developmental stages suggests that ovD19C2 has certain "housekeeping" functions. The elevated expression in the ovary can be interpreted as followed. Since most zygotic genes are not active until 1 hour of development, many gene activities essential for basic cellular process must be provided maternally. A high concentration pool of such gene products may exist in the freshly laid eggs to support embryonic development before the zygotic genome is activated.

The materials in this pool may be consumed progressively until the zygotic genes are active; and the demand and supply of these products reaches a dynamic balance. In order to generate such a high concentration pool, the enhancement and expression of these genes in the ovary is necessary. This can be achieved in two ways. The first way is by gene amplification, since polyploidization in the nurse cells is extensive (see Introduction). The second is by possessing an enhancer which responds to certain cellular factors existing in ovarian cells. In the case of ovD19C2, the high level of expression in testes favours the second argument, because no gene-amplification mechanism has been identified during spermatogenesis. Alternatively, there may be an enhancer element which responds to a cellular factor in both male and female germ-line cells, as in situ hybridization shows that enhanced transcription of ovD19C2 happens in germ-line cells in the ovary and testis.

Given it is a ubiquitin C-terminal hydrolase (5.4.3), ovD19C2 protein may be involved in a variety of cellular processes. During embryonic pattern formation (see Introduction), a number of proteins are required as morphogenic signals to specify certain structures of the embryo. Since these proteins involve very specific functions, and embryo development is very fast, rapid turnover must be necessary. Ubiquitin-mediated protein degradation may provide a way for rapid turnover of these proteins. As described in the introduction of this chapter, 90% of short-lived proteins are degraded by the ubiquitin pathway. Although there is no direct evidence that ubiquitin C-terminal hydrolase participates in protein degradation, it might be required for generating free ubiquitin after the target protein is degraded. Also, the ubiquitin pathway of protein degradation may be used to processing yolk proteins, to provide energy and raw materials for the developing embryo. During early embryonic development, cell proliferation is predominant, and extensive production of ribosomes is required to support protein synthesis. In this process, UbCEPs must be cleaved by certain ubiquitin hydrolases and CEPs are incorporated into ribosome. It is known that yeast YCH1 has the ability to cleave a ubiquitin-protein fusion which contains 72 amino acids in the C-

terminal extension (Miller *et al.*, 1989) . As its counterpart in *Drosophila*, the ovD19C2 protein may possess the same ability. Thus this enzyme may also has a role in ribosome assembly in embryogenesis

The result of in situ hybridization to the ovD19C2 mRNA has provided more information about the expression pattern of ovD19C2. In the ovary, the ovD19C2 probe detected a strong signal in nurse cells. This elevated transcription starts from stage 9 of oogenesis and peaks at stage 10 (Figure 6.2A.B). The ovD19C2 message is also detected in oocyte of stage 10-14, and early embryos (Figure 6.3C). These results are consistent with the results of Northern analysis. The mRNA is also located in degenerating nurse cells during stage 12-14. Perhaps the activity of ovD19C2 is required for the degeneration of these nurse cells, as ubiquitin is involved in protein degradation. But similar phenomena have been observed for other genes (Grimes; Slee; personal communications). The transcript of the *Drosophila* IMPD gene is also detected in degenerating nurse cells; but there is no known function for IMPD in protein degradation. The possible reason for the detection of these transcripts in degenerating nurse cells is that these mRNAs failed to be transported into oocyte. In the male gonads, the hybridization signal is detected in the testes. By comparison, the signal in the accessory glands is very weak (Figure 6.4A.B). The stronger signal at the inner side of the coiling testis indicates that the transcription occurs in the spermatocyte cells. From this point of view, the expression of ovD19C2 is also enhanced in the male germ-line cells. The stronger signals detected in nurse cells and spermatocyte cells suggest that the enhancement of expression of ovD19C2 is limited to germ-line cells. This ultimately leads to the conclusion that this gene contains a regulatory element, which responds to cellular factors only existing in germ-line cells.

CHAPTER SEVEN

PRELIMINARY EXPERIMENTS TOWARDS UNDERSTANDING THE FUNCTION OF OVD19C2

7.1 The Expression of ovD19C2 in E. coli

As described in chapter 5, the predicted protein of ovD19C2 has a extensive similarity to the human ubiquitin carboxyl terminal hydrolase, however, to confirm this experimentally, an enzymatic essay is needed. This requires purification of the protein. In addition, although the Northern analysis and in situ hybridization showed the transcription sites of this gene, where and when its mRNA is translated remains to be determined. For this, we need to obtain antibodies against this protein.

To express the ovD19C2 protein for further studies (raising antibodies and enzymatic analysis), an expression construct was made. The pBD19C2 insert was excised with HincII and EcoRI, and then cloned into the SmaI and EcoRI sites of the vector pGEX-2T (Pharmacia). Once expressed, the ovD19C2 protein should be fused to the carboxyl terminus of glutathione S-transferase (GST) of *Schistosoma japonicum* (Smith *et al.*, 1986). The expression is driven by the tac promoter which is inducible by adding IPTG. The fusion protein can be cleaved with thrombin, at the junction between the two proteins. The resulting construct is pGEX-D19C2, and was used to transform the E. coli strain NM522.

After IPTG induction, the bacterial lysate was then electrophoresed on 10% SDS-PAGE gel. The result is shown in Figure 7.1. Lane 2 is the protein from the parent plasmid. Lanes 2-4 are the GST-ovD19C2 fusion protein induced with different concentrations of IPTG. The molecular weight of the fusion protein form this gel is estimated at 53 kD. After subtracting the glutathione S-transferase, ovD19C2 protein is estimated at 26 kD. This agrees with the molecular weight predicted from the sequence data.

7.2 Excision of the P[lac,ry⁺]A insertion from D19

7.2.1 Mobilizing the P[lac, ry⁺]A transposon by Crossing to P[ry⁺, Δ 2-3](99B)

In Chapter 6, the possible functions of ovD19C2 have been discussed, however, these can only be examined when mutations on this gene are available. This gene maps at 98F on the third

Figure 7.1

SDS-PAGE gel of GST-ovD19C2 fusion protein.

1,8 = molecular weight marker

2-3 = pGEX-2T without ovD19C2

4-7 = GST-ovD19C2 fusion protein

In lane 2 and 3, the parent plasmid pGEX-2T expressed the GST protein of 27 kD. In lane 4-7, the GST-ovD19C2 fusion protein is about 52 kD in molecular weight.

chromosome. A search for existing deficiencies and point mutations in this cytological region was made, and Table 7.1 is a list of genes and chromosome rearrangement in this region. The only deficiency in this region is Df(3)Mg32 which covers 97F to 100C chromosomal region (Lindsley and Zimm, 1985), and there are no relevant point mutations exist in this region. Thus the generation of mutations is an important step towards understanding the function of the gene we have cloned. There are many ways to generate mutations: such as gamma-ray radiation, EMS mutagenesis and P-element mediated mutagenesis. Another way relevant to the line D19 is to mobilize existing P-elements. The imprecise excision of the P-elements can cause deletions of the flanking genomic sequence (Fasano *et al.*, 1991). Thus in D19, excision of the P[*lac*, *ry*⁺]A transposon is a practical approach to generate deletions surrounding the transposon. As GH1 and GH4 are near the transposon, it may be possible to delete these sequences. The deletions also would be very useful for screening point mutations and insertional mutations in this gene to be generated by EMS and P-element mediated mutagenesis.

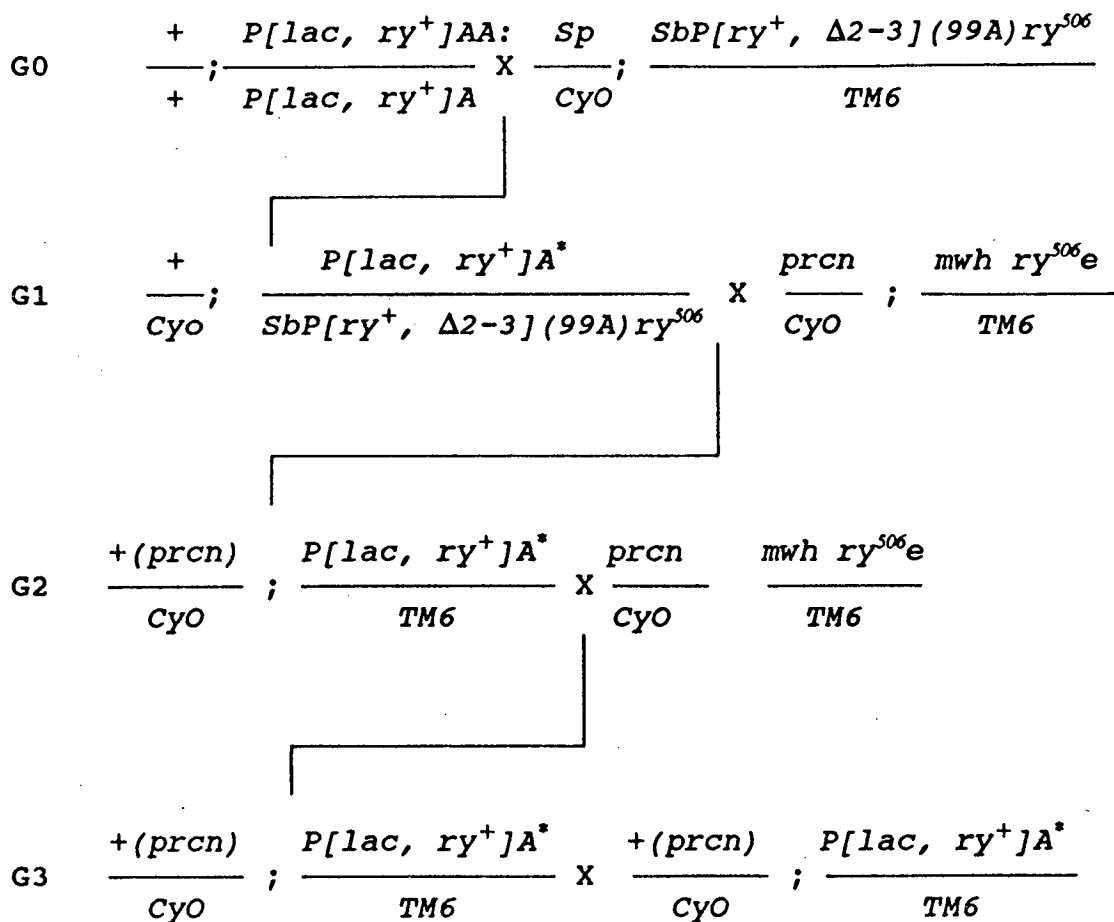
D19 flies carrying the P[*lac*, *ry*⁺]A insertion on 98F of chromosome 3 (5.3.2) were crossed to the E#2 strain, which carries a second P-element on chromosome 3. This element, P[*ry*⁺, $\Delta 2-3$](99B), is not itself mobile, but can provide transposase to mobilize other P-elements in the genome (Robertson *et al.*, 1988). The crossing scheme is shown in Figure 7.2. In G1, males were selected for *Cy*⁻ and *Sb*⁻, which gave a +/*CyO*; P[*lac*, *ry*⁺]A/*Sb*⁻ P[*ry*⁺, $\Delta 2-3$] genotype. These males are called jumpstart males. A single jumpstart male was crossed to 5 virgin females of the *pr*⁻*cn*/*CyO*; *mwh-ry*⁵⁰⁶ *e*⁻/*TM6* genotype, to balance the second and third chromosomes. The G2 progeny of this cross were selected for *Ubx*⁻ and *Sb*⁺, representing P[*lac*, *ry*⁺]A^{*}/*TM6* (here the * mark means the transposon may have been mobilized). Since the target was the third chromosome, no selection was made for chromosome 2. The virgin females of P[*lac*, *ry*⁺]A^{*}/*TM6* were then back-crossed to *pr*⁻*cn*/*CyO*; *mwh-ry*⁵⁰⁶ *e*⁻ males. In G3, flies were selected for *Ubx*⁻ and *e*⁺, and a sibling mating was carried out. In the next

Figure 7.2

The genetic cross scheme for mobilizing the P[lac, ry⁺]A from D19.

Figure 7.2

Scheme of Genetic Crosses of Mobilization of P[lac, ry⁺]A



generation, the criterion for selection was if there were any flies which were homozygous for chromosome 3. If no *Ubx*⁺ flies were present within a certain line, a lethal event must have occurred on chromosome 3. If flies homozygous for chromosome 3 were viable, there are two possible explanations. Firstly, the excision of *P[lac, ry*⁺*]/A* may not lead to a mutation or, secondly, any mutations present must be zygotically non-lethal and recessive. These need to be checked further for fertility and other possible phenotypes.

150 independent lines were obtained in G4, and 100 of these lines were examined for *Ubx* phenotypes. In the 100 lines examined, 46 lines were found where flies all showed the *Ubx*-phenotype, indicating that these lines were homozygous lethal for the third chromosome, and the size of assumed deletions caused by the excision of the *P[lac, ry*⁺*]/A* in these lines is probably large and there is a vital gene located near to the transposon. This is a surprising result, since Sved *et al.* (1990) reported that the frequency of recombination surrounding the P-elements in such a cross is about 0.5 to 1%. The possible reason of such a large number of deletions will be discussed later. In the remaining 54 lines, the population were a mixture of *Ubx*⁺ and *Ubx*-flies. Of these 54 lines, 50 showed *ry*- Phenotype. This indicates that the *P[lac, ry*⁺*]/A* insertion must have been excised in these flies. Another 5 lines were *ry*⁺; either the transposon has not been mobilized, or it has jumped to a new chromosome location. This suggests that there is a very high rate of excision of the P-element and of the creation of deletions in the process of excision.

7.2.2 Did deletion Occur? Analysis of Deletions by Southern Blotting

As a preliminary experiment, DNA from 8 lines producing only *Ubx*-flies was digested with *EcoRI* and a Southern blot carried out. This was probed with GH4 probe, and the result is shown in Figure 7.3. The GH4 probe detected a single 11 kb *EcoRI* fragment in the DNA of all the eight lines, as well as in that of Oregon R. This shows that the assumed deletion did not occur in this fragment within these eight lines. Thus further analysis is required to examine

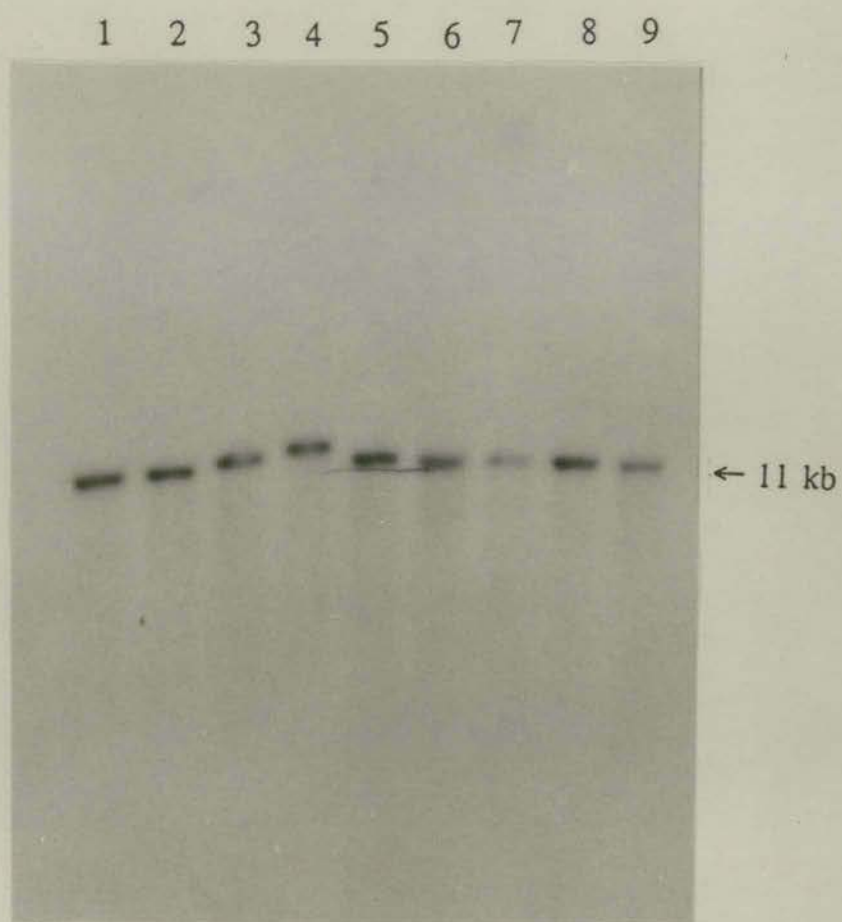
Figure 7.3

Southern blot of EcoRI digested DNA from 8 Ubx⁻ lines generated by excision of the P[lac, ry⁺] insertion from D19 genome. GH4 probe.

1 = Oregon R DNA, EcoRI digestion

**2-8 = DNA from different Ubx⁻ lines,
EcoRI digestion**

FIGURE 7.3



whether these deletion occurred in the region flanking 5' end of the P[*lac, ry⁺*]A insertion, i.e. in the region of the GH3 or GH1, or if they were 3' of the transposon and included the GH2 fragment. The remaining lines also need to be examined to determine if any deletions extend to GH4.

7.3 Discussion

During attempts to generate deletions surrounding the p[*lac, ry⁺*]A insertion, large numbers (46 of 100) of lethal lines were obtained. Sved *et al.* (1990) reported that the rate of recombination in the region of the nonautonomous P-element, P[CaSpeR], is from 0.5% to 1%, after induction by P[*ry⁺, Δ 2-3*](99B). One reason for such a high proportion of lines showing a lethal phenotype may be the special chromosomal location of the transposon in D19. The P[*lac, ry⁺*]A insertion is located at 98F on chromosome 3, which is near to P[*ry⁺, Δ 2-3*](99B). Sved *et al.* (1991) have show that up to 20% recombination can be induced by homologous P-elements.

According to Engles' "cut and paste" model, the transposition of P-elements leaves behind a double-strand gap at the donor site. This gap is then repaired with the sister strand acting as a template. In the homozygotes, the result is the replacement of the P-element at the donor position. If the individual is heterozygous for the P-insertion, and the corresponding chromosome is used as a template, the result is the precise loss of the P-element (Engels, *et al* 1990). In D19, if P[*ry⁺, Δ 2-3*](99B) serves as a template for repairing the gap left by the transposition of P[*lac, ry⁺*]A, it will lead to deletion of the sequence surrounding the P[*lac, ry⁺*] insertion. In addition, in the case of P[*lac, ry⁺*]A, the gap left by excision of this element is beyond 10 kb, according to Sved *et al.* (1991), when the gap is larger than several kilobases, the frequency of crossing-over will be over 50%. If the excision of P[*lac, ry⁺*] occurred after it paired with P[*ry⁺, Δ 2-3*](99B), and a crossing-over happened, it would deleted the sequence between these two elements, i.e. from 98F to 99B. Both Southern analysis and chromosome squashes can be employed to test these hypotheses. In the former case, genomic DNA flanking the P[*lac, ry⁺*] insertion can used to probe Southern

blots made from these *Ubx*⁻ flies. In the latter case loop will form on the wildtype chromatid of the polytene chromosome at the location where deletion has occurred in heterozygotes. If this theory is true, one would expected that will happen in a similar region of DNA.

FINAL DISCUSSION

When O'Kane and Gehring first designed the enhancer trap technique, it was based on the assumption that if the enhancer detector was inserted near a developmentally regulated gene, the enhancer would lead to the expression of the *lacZ* reporter gene with the same pattern as the gene normally controlled by this enhancer (O'Kane and Gehring, 1987). Subsequent experiments using this technique have indeed shown that large numbers of these enhancer trap lines do express the P-*lacZ* fusion gene in tissue specific patterns in the embryo (Bier *et al.*, 1989; Bellen *et al.*, 1989), ovary (Grossniklaus *et al.*, 1989), and in several cases the expression patterns of the *lacZ* fusion gene is similar to that of known genes mapped at the same chromosomal location. Direct evidence comes from a recent report of cloning of a gene called *teashirt* using the enhancer technique (Fasano *et al.*, 1991). In this case, a transposant strain carrying the P[*lac*, *ry*⁺]A insertion showed the β -galactosidase expression in the truck region of the embryo, i.e. posterior labial segment to anterior part of the eighth abdominal segment. In situ hybridization of a probe made from the *teashirt* gene to embryos showed that this gene was expressed in a way similar to the *lacZ* fusion gene. Molecular analysis showed that the transposon was inserted within the 5' untranslated region of the *teashirt* gene.

In this project, 61 transposants carrying the enhancer detector P[*lac*, *ry*⁺] have been examined for the expression patterns of the *lacZ* fusion gene in the adult. Results show that in almost all cases (54/56) the expression of this reporter gene in adult gonads is temporally and spatially regulated. Many lines do show very interesting staining patterns, such as those exhibiting β -galactosidase expression in the germarium or polar cells. (Chapter 3).

A 16.5 kb genomic fragment surrounding the p[*lac*, *ry*⁺]A insertion has been isolated from the line D19. This line expresses the *lacZ* fusion gene in both male and female germ-line cells in the adult (Figure 3.5A,B). Northern analysis shows that the genomic fragment GH1 flanking the 5' end of the transposon detects two ovary specific transcripts. No corresponding cDNAs have yet been

isolated. This is probably because the transcripts are not abundant, so a larger number of recombinant phages should be screened. Alternatively, the cDNA library used may not be representative, since 1.5×10^5 recombinant phages have been screened. This library has been amplified once. The construction of a new ovary cDNA library is under way in this laboratory, and this genomic fragment will be used to screen the new library, and isolate the cDNA corresponding to the transcription units immediately flanking the transposon.

Another genomic fragment (GH4), which is further upstream of the p[*lac*, *ry*⁺] insertion, has been analyzed in detail. This fragment contains a gene encoding a 1.1 kb mRNA. This gene, *ovD19C2*, is differentially expressed during development. It is expressed at low levels during most of the developmental stages, and its expression is enhanced within the adult germ-line during oogenesis and spermatogenesis. This may be a common expression pattern for many "housekeeping" genes which are constitutively expressed during development. During early embryonic development, before zygotic genes are transcribed, these gene products must be provided maternally. This requires high level expression of these genes during oogenesis and the storing of their products in the oocyte. In the case of *ovD19C2*, the high level expression in the male germ-line cells may be due to the fact that the same enhancer is also active in the male germ-line, even though the activity of this gene is not absolutely required for spermatogenesis. This was the case for the maternal-effect gene *vasa* (Lasco and Ashburner, 1990). It will be very interesting to further investigate whether or not the sequences flanking GH4 contain the enhancer which is responsible for the elevated expression of *ovD19C2* in the ovary and testis. It is likely the genomic fragment GH3 which flanks the 5'end of GH4, in respect of the transcriptional direction of *ovD19C2*, contains the enhancer which elevates the transcription of *ovD19C2*. This can be achieved by transforming flies with a transposon, in which a reporter gene (*lacZ*) is put under the control of the GH3 genomic fragments, and examine whether or not this reporter gene shows the germ-line-

specific expression. If there is such an enhancer, the deletion of it must abolish the enhanced expression of *ovD19C2* in the germ-line, and the function of this gene during oogenesis and spermatogenesis can be accessed. Sequencing analysis of *ovD19* and the corresponding genomic DNA shows that this gene encodes a 227 amino-acid protein. The amino acid sequence of the predicted protein shows a high level similarity to the human ubiquitin carboxyl terminal hydrolase (51.1% identity to human UCH-L3, and 47% to human UCH-L1). The *ovD19C2* protein has all the important features defining this family of enzymes (discussed in Chapter 5).

In order to further investigate the function of the *ovD19C2* protein, the cDNA has been cloned into an expression vector, and preliminary experiments show that the expressed protein has the correct molecular weight. This expression construct can be used for expression and purification of this protein, and to demonstrate experimentally whether or not this enzyme has the ability to cleave ubiquitin fusion proteins, or other ubiquitin conjugates, and which kind of ubiquitin conjugates this enzyme is designed specifically to cleave.

Another application of the expression construct is to over-express this protein, and raise an antibody against it. Once the antibody is available, the expression of this gene can then be explored at the protein level.

It is very important to obtain mutations of this gene to study its function and regulation. By mobilizing the *p[lac, ry⁺]*A insertion from the genome of D19, large numbers of zygotic lethal lines have been obtained. It is likely that the lethal phenotype in these lines is caused by some deletions of genomic sequences surrounding the transposon. It is possible that in some of these lines the deletion may reach the GH4 fragment. However, this needs to be further investigated. Since the *P[lac, ry]*A insertion is located very close to a HindIII site of λ D10 (Figure 4.4), it should be relatively easy to test in which direction the deletion occurred by using GH1 and GH2 fragments to probe Southern blots containing DNA from these lethal lines. Using the same technique, a deletion of 40 kb of sequences surrounding the transposon was generated in one transposant

strain (Fasano *et al.*, 1991). Other approaches can also be used to generate deletions surrounding the transposon, for example, by gamma radiation. Because D19 flies also carry the genetic marker *ry*⁺, it is relatively easy to screen deletions surrounding the marker. Recently, a new method has been created to make insertional mutations on targeted genes (Kaiser and Goodwin, 1989; Ballinger and Benser, 1989). This is achieved by crossing the *Drosophila* strain Birm-2, which carries 17 defective P-elements, to the transposase source P[*ry*⁺, Δ 2-3](99B) to mobilize the P-element into the gene of interest, and screening the mutations by polymerase chain reaction with primers one designed from the terminal repeat sequence of the P-element and another from the targeted gene. An application of this technique has been successful (K. Lineruth, personal communication). As *ovD19C2* and corresponding genomic fragment have been sequenced, this experiment can now be conducted.

Once there is a mutation background, this gene can be modified in various ways *in vitro*, and introduced back into the mutation background. Then its function and regulation can be fully accessed. For *ovD19C2*, it is extremely interesting to know how the expression of this gene is enhanced during oogenesis and spermatogenesis.

Several genes involved in ubiquitin pathway have been cloned in *D. melanogaster*, they are the polyubiquitin gene, *UB3-D* (*UbCEP80*) gene (Lee *et al.*, 1988) and the *UbcD1* gene (encoding E1 enzyme) (Treiver *et al.*, 1992). However, little is known about their functions and regulations during development. In this project, the results have shown that the putative *Drosophila* ubiquitin carboxyl terminal hydrolase gene is expressed differentially during development, in which its expression is greatly enhanced during oogenesis. This implies that other genes involved in the ubiquitin pathway may also be expressed in a similar way. In regard of the functions of ubiquitination (discussed in Chapter Six), it is very likely that ubiquitination and deubiquitination are very much involved in embryogenesis. Firstly, the C-terminal extension proteins encoded by the *UbCEP* genes are ribosome proteins (Finley

et al., 1989), and during early embryonic development, there is an extensive synthesis of protein which requires large amounts of ribosomes. Secondly, about 90% of short lived protein are degraded by the ubiquitin mediated proteolysis, it is possible the products of those genes involved in the embryonic pattern formation are subjected to degradation in the ubiquitin pathway, since their function is usually only required for defining specific domains of the embryo, and only for a short time. Further analysis of these genes and cloning new genes involved in the ubiquitin pathway in *D. melanogaster* will improve our understanding of how the ubiquitin system functions in this organism.

Compared to embryogenesis, little is known about ovary development. The further characterization of these lines, especially those showing the *lacZ* staining during early stages of oogenesis such as germarium and polar cells, and cloning and analyzing these genes specifically active in such cells will no doubt provide a better understanding of the complex process of oogenesis. These genes may have decision making roles in cell determination, differentiation and pattern formation during the development of the ovary.

REFERENCES

- Anderson, K.V and Nüsslein-Volhard, C. (1986) Dorsal-group genes in *Drosophila*. In *Gametogenesis and the Early Embryo* (Ed. J. Gall), Symp. Soc. devl Biol. 43, 177-194.
- Anderson, K.V., Bokla, L. and Nüsslein-Volhard, C. (1985b) *Cell* 42, 791-798.
- Anderson, K.V., Jürgens, G. and Nüsslein-Volhard, C. (1985a) *Cell* 42, 779-789.
- Appleyard, R.K. (1954) *Genetics* 39, 440
- Arribas, C., Sampedro, J. and Izquierdo, M. (1986) *Biochem-Biophys. Acta* 868, 119-127.
- Bakken, A.H. (1973) *Devl. Biol.* 33, 100-122.
- Ballinger, D.G. and Benzer, B. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 9402-9406.
- Barnett, T., Pacht, C., Gergen, J.P. and Wensink, P.C. (1980) *Cell* 21, 729-738.
- Bellen, H.J., O'Kane, C.J., Wilson, C., Grossniklaus, U., Kurth Pearson, R. and Gehring, W.J. (1989) *Genes Devl.* 3, 1288-1300.
- Belote, J.M., Jandler, A.M., Wolfner, M.F., Linvak, K.L. and Baker, B.S. (1985) *Cell* 40, 339-348.
- Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Browder, Z. and Buck, G. (1988) *EMBO J.* 7, 1121-1127.
- Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., Ackerman, L., Carretto, R., Uemara, T., Grell, E., Jan, L.Y. and Jan, Y.N. (1989) *Genes Devl.* 3, 1273-1287.

- Blow, D.M., Birktoft, J.J. and Hartley, B.S. (1969) *Nature* 221, 337-340.
- Bond, U. and Schlesinger, M.J. (1985) *Mol. Cell Biol.* 5, 949-959.
- Brandbury, E.M. (1985) *J. Biol. Chem.* 260, 5147-5153.
- Brink, N.G. (1968) *Mutation Res.* 5, 192-194.
- Bownes, M. (1990) *Mol. Gen. Genet.* 222, 457-460.
- Bownes, M. and Hames, (1978) *FEBS Lett.* 96, 327-330.
- Bownes, M. and Dale, L. (1982). In: *A Handbook of Drosophila Development* (Ed. R. Ransom), Elsevier Biomedical Press, pp. 31-48
- Bownes, M., Scott, A. and Blair, M. (1987) *Development* 101, 931-941.
- Bownes, M. Scott, A. and Shirras, A. (1987) *Development* 103, 119-128
- Brennan, M.D., Weiner, A.J., Goralski, T.J. and Mahowald, A.P. 1982) *Devl. BioL.* 89, 225-236.
- Brower, D.L., Smith, R.J. and Wilcox, M. (1981) *J. Embryol. Exp. Morphol.* 63, 233-242.
- Brown, E.H. and King, R.C. (1964) *Growth* 28, 41-81.

- Burtis, K.C., Coschigano, K.T. Baker, B.S. and Wensink, P.C. (1991) EMBO J. 10, 2577-2582.
- Chung, C.T., Niemela, S.L. and Miller, R.H. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 2172-2175.
- Cooly, L., Kelly, R. and Spradling, A. (1988) Science 239, 1121-1128.
- Cooley, L., Xue, F. and Verheyen, E. (1990). Information listed in Abstracts from *Drosophila* Meeting, Crete, 1990.
- Day, I.J. and Thompson, R.J. (1987) FEBS Lett. 210, 157-160.
- Day, J.W. and Gell, R.F. (1976) Genetics 83, 67-79.
- Digan, M.E., Spradling, A.C., Waring, G.L. and Mahowald, A.P. (1979) The Genetic Analysis of Chorion Morphogenesis in *Drosophila*. In: Eukaryotic Gene Regulation, ICN-UCLA Symp. (ED. Axel, A. Maniatis, T. and Fox, C.F.), Vol.23. Academic Press, New York.
- Doyle, H., Kraut, R. and Levine, M. (1989) Genes Devl. 3, 1518-1533.
- Driever, W. and Nüsslein-Volhard, C. (1988a) Cell 54, 83-94.
- Driever, W. and Nüsslein-Volhard, C. (1988b) Cell 54, 95-104.
- Dworkin-Rastl, E., Shrutkowiski, A. and Dworkin, M.B. (1984) Cell 39, 321-325.

Edgar, B.A. and Schubiger, G. (1986) *Cell* 44, 871-877.

Engels, W.R., Johnson-Schlitz, M.D., Eggleston, W.B. and Sved, J.A. (1990) *Cell* 62, 515-525.

Fasano, L. and Kerridge, S. (1988) *Development* 104, 245-253.

Fasano, L., Röger L., Core, N., Alexandre, E., Vola, C., Jacq, B. and Kerridge, S. (1991) *Cell* 64, 63-79.

Feinberg, A.P. and Vogelstein, B (1983) *Anal. Biochem.* 132, 6-13.

Feinberg, A.P. and Vogelstein, B (1984) *Anal. Biochem.* 137, 226-267

Finley, D. and Chua, V. (1991) *Annu. Rev. Cell Biol.* 7, 25-69.

Finley, D., Ciechanover, A. and Varshavsky, A. (1984) *Cell* 37, 434-455.

Finley, D., Ozakaynak, E. and Varshavsky, A. (1987) *Cell* 48, 1035-1046.

Finley, D., Barte, B. and Varshavsky, A. (1989) *Nature* 338, 394-401.

- Frigerio, G., Noll, M. and Nüsslein-Volhard, C. (1988) EMBO J 7, 1749-1756.
- Frohnhofer, H.G. and Nüsslein-Volhard, C. (1986) Nature 324, 120-125.
- Frohnhofer *et al* 1986, J.Embryol. exp. Morphol. 97 suppl. 169-179.
- Gans, M., Audit, C. and Masson, M. (1975) Genetics 81, 883-704.
- Garica-Bellido, A. and Rubbins, L.G. (1983) Genetics 103, 235.
- Garabedian, M.J., Hung, M-C. and Wensink, P.C. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 1396-1400.
- Garabedian, M.J., Shepherd, B.M. and Wensink, P.C. (1986) Cell 45, 859-867.
- Gaul, V. and Jäckle, H. (1987) Cell 51, 549-555.
- Gausung, K. and Barkardottir, R. (1986) Eur. J. Biochem. 158, 57-62.
- Gerttula, S., Jin, Y. and Anderson, K.V. (1988) Genetics 119, 123-133.
- Ghosh, S., Gifford, A.M., Driever, L.R., Tempst, P., Nolan, G.P. and Baltimore, D. (1990) Cell 62, 1019-1029.
- Gough, J.A. and Murray, N. (1983) J. Mol. Biol. 166, 1-19.

Grantham, R., Gautier, C., Gouy, M., Mercier, R. and Pave, A. (1980) Nucl. Acids Res. 8, r49-62.

Griffin-Shea, R., Thiresos and Kafatos, F.C. (1982) Devl. Biol. 91, 325-336.

Grossniklaus, U., Bellen, H.J., Willson, C. and Gehring, W. Gertlula, S., Jin, Y. and Anderson, K.V. (1988) Genetics, 119, 123-133.

Haas, A.L. and Bright, P.M. (1988) J. Biol. Chem. 263, 13258-13267.

Haenlin, M., Roos, C., Cassab, A. and Mohier, E. (1987) EMBO J. 6, 801-807.

Hall, J.C. (1972). Genetics 71, 367-400.

Hashimoto, C., Hudson, K.L. and Anderson, K.V. (1988). Cell 52, 269-279.

Hay, B., Jan, L.Y. and Jan, Y.N. (1988). Cell 55, 577-587

Henning, W. (1967) Chromosoma 22, 294-357.

Holzwarth, K.W., Spectar, C. and Gottleib, F.J. (1974). Wilhelm Roux's Arch. Devl. Biol. 174, 260-275.

Hough-Evan, B.R., Jacobs,-Lorena, M., Cumming, M.R., Britten, R.J. and Davidson, E.H. (1980) *Genetics* 95, 81-94.

Huang, S.Y., Barnard, M.B., Xu, M., Matsui, S.I., Rose, S.M. and Garrard, W.T. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 3738-3742.

Hung ,M. and Wensink, P.C. (1983) *J. Mol. Biol.* 164, 481-489.

Huynh, T., Young, R.A. and Davis, R. (1985) In: *DNA Cloning*, Vol. 1, (Ed. Glover, D.), IRL Press, Oxford, Washington, DC, pp56-110.

Illmensee, K. and Mahowald, A.P. (1976) *Exp. Cell Res.* 97,127-140.

Irish, V., Lehmann, R. and Akam, M. (1989) *Nature* 338, 646-648.

Jäckle, H., Gaul, U. and Redemann, N. (1988) *Development* 104 Suppl. 29-34.

Jendrisak, J., Young, R.A. and Engel, J. (1978) In: *Guide to Molecular Cloning Techniques*. (Ed. Berger, S. and Kimmel, A.) Academic Press, New York, pp 359-371.

Jonnalagadda, S., Butt, T.R., Monia, B.P., Mirabelli, C.K. and Gotlib, L. (1989) *J. Biol. Chem.* 264, 10634-10642.

Jowett, T. (1986) Preparation of Nucleic Acids. In: *Drosophila: a practical approach*. (Ed. Roberet, D.B.), IRL Press, Oxford, Washington DC, p278.

Kaiser, K. and Goodwin, S.F. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 87, 1686-1690.

Kambysellis, M.P. (1977) *Amer. Zool.* 17, 535-550.

Karess, R.E. and Rubin, G.M. (1984) *Cell* 38, 135-146.

Keller, E.B. and Noon, W.A. (1985) *Nucl. Acids Res.* 13, 471-481.

Kidwell, M.G. (1979) *Genet. Res. Camb.* 33, 205-217.

Kidwell, M.G., Kidwell, J.F. and Sved, J.A. (1977) *Genetics* 36, 813-833.

Kieran, M., Blank, V., Logeat, F., Vandekerckhove, J., Lottspeich, F., LeBau, O., Urban, M.B., Kourilsky, P., Baeuerle, P.A. and Israel, A. (1990) *Cell* 62, 1007-1018.

King, R.C. (1970) *Ovarian Development in Drosophila melanogaster*. Academic Press, New York.

King, R.C. and Burnett, R.G. (1959) *Science* 129, 1674-1675.

King, R.C. and Mohler, J.D. (1975). The Genetic Analysis of Oogenesis in *Drosophila melanogaster*. In: A Handbook of Genetics (Ed. R.C.King), pp. 757-791. Plenum Press. New York and London.

King, R.C., Robinson, A.C. and Smith, R.F. (1956) Growth 20, 121-157.

Klingerler, M., Erdélyi, M., Szabad, J. and Nüsslein-Volhard, C. (1988) Nature 335, 275-277.

Koch, E.A., Smith, P.A. and King, R.C. (1967) J. Morphol. 121, 55-70.

Koch, E.A. and King, R.C. (1969) Z. Zellforsch. 102, 129-152.

Komitopoulou, K., Gans, M., Margaritis, L.M., Kafatos, F.C. and Levinger, L. and Varshavsky, A. (1982) cell 28, 375-385.

Lasko, P.F. and Ashburner, M. (1990). Genes and Development 4: 905-921

Lee, H., Simon, J.A. and Lis J. (1988) Mol. Cell Biol. 8, 4727-4735.

Lehmann, R. (1988) Development Suppl. 104, 17-27.

Lehmann, R. and Nüsslein-Volhard, C. (1986) Cell 47, 141-152.

- Lehmann, R. and Nüsslein-Volhard, C. (1991) *Development* 112, 679-693.
- Lindsley, D.L. and Grell, E.H. (1968). *Genetic Variations of D.melanogaster*. Carnegie Institute of Washington, Publication No. 627
- Mahowald, A.P. (1962) *J. Exp. Zool.* 151, 201-215.
- Mahowald, A.P. (1971a) *J. Exp. Zool.* 176, 329-344.
- Mahowald, A.P. (1971b) *J. Exp. Zool.* 176, 345-352.
- Mahowald, A.P. and Strasshein, J.M. (1970) *J. Cell Biol.* 45, 306-320.
- Mahowald, A.P. and Tiefert, M. (1970) *Wilhelm Roux Arch.* 165, 8-25.
- Margaritis, L.H. (1985) *Structure and physiology of the eggshell*. In: *Comprehensive insect physiology, biochemistry and pharmacology* (Ed. G.A. Kertut and L.I. Gilbert), Vol. 1, pp113-115. Pergamon Press, Elmsford.
- Margaritis, L.H., Kafatos, F.C. and Petri, W.H. (1980) *J. Cell Sci.* 43, 1-35.
- Masson, M. (1983) *Genetics* 105, 897-920.

- Mayer, A.N. and Wilkinson, K.D. (1989) *Biochemistry* 28, 166-172.
- Meuller, R.D., Yashuda, H., Hatch, C.L., Bonner, W.M. and Monia, B.P., Ecker, D.J., Finley, D. and Crooke, S.T. (1990b) *J. Biol. Chem.* 265, 19356-19361.
- Mével-Ninio, M, Mariol, M-C. and Gans, M. (1989) *EMBO J* 8, 1549-1558
- Miller, H.I., Henzel, W.J., Ridgeway, J.B., Kuang, W.-J. and Chrisholm, V. (1989) *Bio/technology* 7, 698-704.
- Mohan, J. and Ritossa, F.M. (1970). *Devl. Biol.* 22, 495-512.
- Mohler, J.D. (1977) *Genetics* 85, 259-272.
- Mohler, J.D. and Wieschaus, E.F. (1986). *Genetics* 112, 803-822.
- Monia, B.P., Ecker, D.J., Jonnalagadda, S., Marsh, J., Gotlib, L. Butt, T.R. and Crooke, S.T. (1989) *J. biol. Chem.* 264, 4093-4108.
- Murray, N.E. (1983) *Phage lambda and molecular cloning*. In: *Lambda II* (Ed. Hendrix, Stahl, Roberts and Weisberg), Cold Spring Harbor, New York.
- Nickel. B.E., Allis, C.D. and Davie, J.R. (1989) *Biochemistry*, 28, 958-963.

Neurath, H. (1984) *Science* **224**, 350-357.

Nüsslein-Volhard, C. (1991) *Development Suppl.* **1**, 1-10.

Nüsslein-Volhard, C., Frohnhofer, H.J. and Lehmann, R. (1987) *Science* **238**, 1675-1681.

O'Hare, K. and Rubin, G.M. (1983) *Cell* **34**, 25-35.

Ohmachi, T., Giorda, R., Shaw, D.R. and Ennis, H.L. (1989) *Biochemistry* **28**, 5226-5231.

Okada, M. and Koayashi, S. (1987) *Devl. Growth Differ.* **29**, 185-192.

O'Kane, C.J. and Gehring, W.J. (1987) *Proc. Natn. Acad. Sci. U.S.A.* **84**, 9123-9127.

Ozkaynak, E., Finley, D., Solomon, M-J. and Varshavsky, A. (1987) *EMBO J* **6**, 1429-1439.

Parks, S. and Spradling, A.C. (1987) *Genes Devl.* **1**, 497-509.

Parks, S., Wakimoto, and Spradling, A.C. (1986) *Devl. Biol.* **117**, 294-350.

Perrimon, N. and Gans, M. (1983) *Devl. Biol.* **100**, 365-373.

Pauli, D. and Mohoward, A.P. (1990) Trends Genet. 6, 259-264.

Perrimon, N., Engstrom, L. and Mahowald, A.P. (1984) Devl. Biol. 105, 404-414.

Perrimon, N., Mohler, D., Engstrom, L. and Mahowald, A.P. (1986) Genetics 113, 695-712.

Perrimon, N., Engstrom, L. and Mahowald, A.P. (1989) Genetics 121, 333-352.

Pichart, C.M. and Rose, I.A. (1985) J. Biol. Chem. 260, 1573-1581.

Prost, E., Deryckere, F., Roos, C., Haenlin, M., Pantesco, V. and Mohier, E. (1988) Gene & Dev. 2, 891-900.

Postlethwait, J.H., Handler, A.M. and Gray, P.W. (1976) A Genetic Approach to the Study of Juvenile Hormone Control of Vitellogenesis in *Drosophila melanogaster*. In: "The Juvenile Hormones". (L. I. Gilbert, Ed.) Plenum Press. New York.

Postlethwait, J.H. and Jowett, T. (1980) Cell, 20, 671-678.

Price, J.V., Clifford, R.J. and Schupbach, T. (1989). Cell 56, 1085-1092.

- Robertson, H.M., Preston, C.R., Phillis, R.W., Johnson-Schlitz, D.M., Benz, W.K. and Engels, W.R. (1988) *Genetics* 118, 461-470.
- Roth, S., Stein, D. and Nüsslein-Volhard, C. (1989) *Cell* 59, 1189-1202.
- Rubin, G.M., Kigwell, M.G. and Bingham, P.M. (1982) *Cell* 29, 987-944.
- Rubin, G.M. and Spradling, A.C. (1982) *Science* 218, 218, 348-353.
- Rushlow, C.A., Frasch, M., Doyle, H. and Levine, M. (1987) *Nature* 330, 583-586.
- Rushlow, C.A., Han, K., Manley, J.L. and Levine, M. (1989) *Cell* 59, 1165-1177.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
- Santamaria, P. and Nüsslein-Volhard, C. (1983) *EMBO J.* 2, 1695-1699.
- Saunders, R.D.C. and Bownes, M. (1986). *Mol. Gen. Genet.* 205, 577-580.
- Schlesinger, M. and Bond. U. (1987) *Oxf. Surv. Euk. Genes* 4, 77-91.

Schüpbach, T. and Wieschaus, E. (1986a) Wilhelm Roux's Arch. Devl. Biol. 195, 302-317.

Schüpbach, T. and Wieschaus, E. (1986b) Devl. Biol. 113, 443-448.

Schubach, T. and Wieschaus, E. (1989) Genetics 121, 101-117.

Shirras, A.D. and Bownes, M. (1987). Mol. Gen. Genet. 210, 153-155

Sonnenblick,(1941) Proc. Natn. Acad. Sci. U.S.A. 27, 484-489.

Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.

Spradling, A.C. and Mahowald, A.P. (1979) Cell 16, 589-598.

Spradling, A.C. and Rubin, G.M. (1982) Science 218, 341-347.

Sprenger, F. Stevens, L.M. and Nüsslein-Volhard, C. (1989) Nature 338, 478-483.

Steinmann-Zwicky, M. Schmid, H. and Nöthiger, R. (1989) Cell 57, 157-166.

Steinmann-Zwicky, M. (1990) EMBO J 7, 3889-3898.

- Steel, L.F., Smyth, A. and Jacobson, A. (1987) Nucl. Acids Res. 15, 10285-1-298.
- Stevens, L.M., Frohnhofer, H.G., Klingler, M. and Nüsslein-Volhard, C. (1990) Nature 346, 660-663.
- Steward, R. (1989) Cell 59, 1179-1188.
- Storto, P.D. and King, R.C. (1988) Devl. Genet. 9, 91-120.
- Sved, J.A., Eggleston, W.B. and Engels, W.R. (1990) Genetics 124, 331-337.
- Sved, J.A., Blackman, L.M., Gilchrist, A.S. and Engels, W.R. (1991) Mol. Gen Genet. 225, 433-447.
- Swindle, J., Ajika, J., Eisen, H., Sanwal, B., Jacquemot, C.,
- Rubin, G.M. and Spradling, A.C. (1982) Science 218, 348-353.
- Tautz, O. (1988) Nature 332, 281-284.
- Tautz, O. and Pfeifle, C. (1989) Chromosoma 98, 81-85.
- Tearle, R. and Nüsslein-Volhard, C. (1987). *Drosophila* Information Service Volume 66, 209-269.

Thierry-Mieg, D., Masson, M. and Gans M. (1972). C.R. Acad. Paris 275, 2751-2754.

Thisse, B., Stoetzel, C., Gorostiza-Thisse, C. and Perrin-Schmitt, F. (1988) EMBO J. 7, 2175-2183.

Thomas, P.S. (1980). Proc. Natl. Acad. Sci. U.S.A. 77, 5201-

Tobias, J.W. Varshavsky, A. (1991) J. Biol. Chem. 266, 12021-12028.

Treiver, M., Seufert, W. and Jentsch, S. (1992) EMBO J. 11, 367-372.

Wharton, R.P. and Struhl, G. (1989). Cell 59, 881-892.

Wharton, R.P. and Struhl, G. (1990). Information listed in Abstracts from *Drosophila* Meeting, Crete, 1990.

Wiborg, O., Rederson, M.S., Wind, A., Berglund, L.E., Marcker, K.A. and Vuust, J. (1985) EMBO J. 7, 1121-1127.

Wieschaus, E. and Szabad, J. (1979) Devl. Biol. 68, 29-46..

Wieschaus, E., Marsh, J.L. and Gehring, W. (1978) Wilhelm Roux's Arch. devl. Biol. 184, 75-82.

Wilkinson, K.D., Lee, K., Deshpande, S., Duerksen-Hughes, P. and Boss, J.M. (1989) *Science* 246, 670-673.

Wilson, C., Kurth Pearson, R., Bellen, H.J., O'Kane, C.J., Grossniklaus, U. and Gehring, W.J. (1989) *Genes Devl.* 3, 1301-1313.

Young, R.A. and Davis, R.W. (1983) *Science* 222, 778-782

Zalokar, M. (1965) *Rev. Suisse Zool.* 72, 241-261.